

**PCT**WORLD INTELLECTUAL PROPERTY ORGANIZATION  
International Bureau

D3

## INTERNATIONAL APPLICATION PUBLISHED UNDER THE PATENT COOPERATION TREATY (PCT)

|   |           |   |
|---|-----------|---|
| <b>(51) International Patent Classification <sup>6</sup>:</b><br>C07K 2/00, 14/47, 16/18, C12N 5/10,<br>15/09, 15/11, 15/12, 15/63, 15/70, 15/74,<br>15/79  | <b>A1</b> | <b>(11) International Publication Number:</b> WO 97/25341<br><b>(43) International Publication Date:</b> 17 July 1997 (17.07.97)  |
| <b>(21) International Application Number:</b> PCT/US97/00270<br><b>(22) International Filing Date:</b> 7 January 1997 (07.01.97)<br><br><b>(30) Priority Data:</b><br>PN 7450 8 January 1996 (08.01.96) AU<br><br><b>(71) Applicants (for all designated States except US):</b> ST. VINCENT'S INSTITUTE OF MEDICAL RESEARCH [AU/AU]; 41 Victoria Parade, Fitzroy, VIC 3065 (AU). TRUSTEES OF DARTMOUTH COLLEGE [US/US]; Office of Technology Transfer, 11 Rope Ferry Road, Hanover, NH 03755 (US).<br><br><b>(72) Inventors; and</b><br><b>(75) Inventors/Applicants (for US only):</b> KEMP, Bruce, E. [AU/AU]; 20 Kellett Grove, Kew, VIC 3101 (AU). STAPLETON, David, I. [US/AU]; 12 Mint Street, Wantima, VIC 3183 (AU). MITCHELHILL, Kenneth, I. [AU/AU]; 5/14 Westbury Street, East St. Kilda, VIC 3183 (AU). WITTERS, Lee, A. [US/US]; 1399 New Boston Road, Norwich, VT 05055 (US). |           | <b>(74) Agents:</b> LICATA, Jane, Massey et al.; Law Offices of Jane Massey Licata, Suite 201, 210 Lake Drive East, Cherry Hill, NJ 08002 (US).<br><br><b>(81) Designated States:</b> AU, CA, JP, US, European patent (AT, BE, CH, DE, DK, ES, FI, FR, GB, GR, IE, IT, LU, MC, NL, PT, SE).<br><br><b>Published</b><br><i>With international search report.</i> |
| <b>(54) Title:</b> NOVEL AMP ACTIVATED PROTEIN KINASE<br><br><b>(57) Abstract</b><br><br>Polynucleotides of AMPK- $\alpha_1$ , AMPK $\beta$ and AMPK $\gamma$ and polypeptides and biologically active fragments encoded thereby are provided. Vectors and host cells containing these polynucleotides are also provided. In addition, methods of preparing polypeptides and antibodies targeted against these polypeptides are provided.   |           |   |

**FOR THE PURPOSES OF INFORMATION ONLY**

Codes used to identify States party to the PCT on the front pages of pamphlets publishing international applications under the PCT.

|    |                          |    |  |    |                          |
|----|--------------------------|----|--|----|--------------------------|
| AM | Armenia                  | GB | United Kingdom                           | MW | Malawi                   |
| AT | Austria                  | GE | Georgia                                  | MX | Mexico                   |
| AU | Australia                | GN | Guinea                                   | NE | Niger                    |
| BB | Barbados                 | GR | Greece                                   | NL | Netherlands              |
| BE | Belgium                  | HU | Hungary                                  | NO | Norway                   |
| BF | Burkina Faso             | IE | Ireland                                  | NZ | New Zealand              |
| BG | Bulgaria                 | IT | Italy                                    | PL | Poland                   |
| BJ | Benin                    | JP | Japan                                    | PT | Portugal                 |
| BR | Brazil                   | KE | Kenya                                    | RO | Romania                  |
| BY | Belarus                  | KG | Kyrgyzstan                               | RU | Russian Federation       |
| CA | Canada                   | KP | Democratic People's Republic<br>of Korea | SD | Sudan                    |
| CF | Central African Republic | KR | Republic of Korea                        | SE | Sweden                   |
| CG | Congo                    | KZ | Kazakhstan                               | SG | Singapore                |
| CH | Switzerland              | LI | Liechtenstein                            | SI | Slovenia                 |
| CI | Côte d'Ivoire            | LK | Sri Lanka                                | SK | Slovakia                 |
| CM | Cameroon                 | LR | Liberia                                  | SN | Senegal                  |
| CN | China                    | LT | Lithuania                                | SZ | Swaziland                |
| CS | Czechoslovakia           | LU | Luxembourg                               | TD | Chad                     |
| CZ | Czech Republic           | LV | Latvia                                   | TG | Togo                     |
| DE | Germany                  | MC | Monaco                                   | TJ | Tajikistan               |
| DK | Denmark                  | MD | Republic of Moldova                      | TT | Trinidad and Tobago      |
| EE | Estonia                  | MG | Madagascar                               | UA | Ukraine                  |
| ES | Spain                    | ML | Mali                                     | UG | Uganda                   |
| FI | Finland                  | MN | Mongolia                                 | US | United States of America |
| FR | France                   | MR | Mauritania                               | UZ | Uzbekistan               |
| GA | Gabon                    |    |  | VN | Viet Nam                 |

## NOVEL AMP ACTIVATED PROTEIN KINASE

Background of the Invention

The present invention relates to novel AMP protein kinase subunits, to polynucleotides encoding these subunit  
5 proteins and to antibodies which bind to these subunits.

The 5'-AMP-activated protein kinase (AMPK) was initially identified as a protein kinase regulating HMG-CoA reductase (Ferrer et al. (1985) *Biochem. Biophys. Res. Commun.* 132, 497-504). Subsequently, the AMPK was shown to  
10 phosphorylate hepatic acetyl-CoA carboxylase (Carling et al. (1987) *FEBS Lett.* 223, 217-222) and adipose hormone-sensitive lipase (Garton et al. (1989) *Eur. J. Biochem.* 179, 249-254). The AMPK is therefore thought to play a key regulatory role in the synthesis of fatty acids and cholesterol.

15 The AMPK is believed to act as a metabolic stress-sensing protein kinase switching off biosynthetic pathways when cellular ATP levels are depleted and when 5'-AMP rises in response to fuel limitation and/or hypoxia (Corton et al. (1994) *Current Biology* 4, 315-324). Partial amino acid  
20 sequencing of hepatic AMPK (Mitchelhill et al. (1994) *J. Biol. Chem.* 269, 2361-2364; Stapleton et al. (1994) *J. Biol. Chem.* 269, 29343-29346) revealed that it consists of 3 subunits, the catalytic subunit  $\alpha$  (63 kDa), and two non-catalytic subunits,  $\beta$  (40 kDa) and  $\gamma$  (38 kDa).

25 The AMPK is a member of the yeast SNF1 protein kinase subfamily that includes protein kinases present in plants, nematodes and humans. The AMPK catalytic subunit,  $\alpha$ , has a

strong sequence identity to the catalytic domain of the yeast protein kinase SNF1, which is involved in the induction of invertase (SUC2) under conditions of nutritional stress due to glucose starvation (Celenza, J.L. and Carlson, M. (1986) *Science* 233, 1175-1180). Both snf1p and the AMPK require phosphorylation by an activating protein kinase for full activity. The sequence relationship between snf1p and AMPK led to the finding that these enzymes share functional similarities, both phosphorylating and inactivating yeast acetyl-CoA carboxylase (Woods et al. (1994) *J. Biol. Chem.* 269, 19509-19516; Witters, L.A. and Watts, T.D. (1990) *Biochem. Biophys. Res. Commun.* 169, 369-376). The non-catalytic  $\beta$  and  $\gamma$  subunits of AMPK are also related to proteins that interact with snf1p; the  $\beta$  subunit is related to the SIP1/ SIP2 /GAL83 family of transcription regulators and the  $\gamma$  subunit to SNF4 (also called CAT-3) (Yang et al. (1994) *EMBO J.* 13, 5878-5886).

An isoform of the mammalian AMPK catalytic subunit has previously been cloned (Carling et al. (1994) *J. Biol. Chem.* 269, 11442-11448) and is referred to herein as AMPK  $\alpha_2$ . The sequence of AMPK is disclosed in WO 94/28116. The AMPK  $\alpha_2$  does not complement SNF1 in yeast, indicating that their full range of functions are not identical.

A novel isoform of the mammalian AMPK catalytic subunit has now been identified and is referred to herein as AMPK  $\alpha_1$ . In addition, full-length cDNAs for the mammalian AMPK  $\beta$  and AMPK  $\gamma$  subunits have now been cloned and polypeptides encoded thereby purified.

#### Summary of the Invention

Accordingly, a first aspect of the present invention provides an isolated polynucleotide which encodes mammalian AMPK  $\alpha_1$  or a sequence which hybridizes thereto with the proviso that the sequence does not hybridize to mammalian AMPK  $\alpha_2$  as defined in Table 1 or Table 5 of WO 94/28116. In a preferred embodiment, the polynucleotide comprises SEQ ID NO: 44. Also provided are vectors comprising such a polynucleotide, a host

cell transformed with such a vector and recombinant proteins encoded by such a polynucleotide.

In a second aspect, the present invention provides a method of producing mammalian AMPK  $\alpha_1$  which comprises culturing  
5 the cell including the polynucleotide of the first aspect of the present invention under conditions which allow expression of the polynucleotide encoding AMPK  $\alpha_1$  and recovering the expressed AMPK  $\alpha_1$ .

In a third aspect, the present invention provides an  
10 oligonucleotide probe of at least 10 nucleotides, the oligonucleotide probe having a sequence such that the probe hybridizes selectively to the polynucleotide of the first aspect of the present invention. By "hybridizes selectively" it is meant that the probe does not hybridize to a  
15 polynucleotide encoding mammalian AMPK  $\alpha_2$  as defined in Table 1 or Table 5 of WO 94/28116. The oligonucleotide probe may include at least about 5 contiguous nucleotides from the polynucleotide sequence which encodes amino acids 352-366. It will be understood by those of skill in the art that the  
20 oligonucleotide probes according to the third aspect of this invention may be used in a number of procedures. These include the analysis of gene regulatory elements; the analysis of gene expression *in vivo*; and the identification of homologous mammalian and non-mammalian cDNAs including the associated  
25 kinase-kinase.

In a fourth aspect, the present invention provides a substantially purified polypeptide encoded by a polynucleotide of the present invention or a biologically active fragment thereof with the proviso that the fragment is not present in  
30 mammalian AMPK  $\alpha_2$  as defined in Figure 3A of WO 94/28116. In a preferred embodiment, the purified polypeptide comprises at least a portion of SEQ ID NOs: 1-43. Also preferred are biologically active fragments comprising at least 8 contiguous amino acids from the sequence DFYLATSPDPSFLDDHHLTR (SEQ ID NO:  
35 45). By "biologically active fragment" it is meant a fragment which retains at least one of the activities of native AMPK  $\alpha_1$  which activities include (i) the ability to stimulate

phosphorylation of protein molecules; and (ii) the ability to mimic the binding of native AMPK  $\alpha_1$  to at least one antibody or ligand molecule.

It will be appreciated by those skilled in the art that a number of modifications may be made to the polypeptides and fragments of the present invention without deleteriously effecting the biological activity of the polypeptides or fragments. This may be achieved by various changes, such as sulfation, phosphorylation, nitration and halogenation; or by amino acid insertions, deletions and substitutions, either conservative or non-conservative (e.g., D-amino acids, desamino acids) in the peptide sequence where such changes do not substantially alter the overall biological activity of the peptide. By conservative substitutions the intended combinations are: G,A; V,I,L,M; D,E; N,Q; S,T; K,R,H; F,Y,W,H; and P, N $\alpha$ -alkylamino acids.

It is also possible to add various groups to the polypeptides or fragments of the present invention to confer advantages such as increased potency of extended half-life in vivo, without substantially altering the overall biological activity of the peptide.

The mammalian AMPK  $\alpha_1$  polypeptide of the present invention may be used to identify compounds which regulate the action of this kinase. Such compounds are desirable since, for example, they may be used to reduce the biosynthesis of cholesterol and fatty acids. They may also be used to inhibit the release of these from intracellular stores by HSL. In addition, they may be used to reduce cellular malonyl CoA levels and promote the  $\beta$ -oxidation of fatty acids by the mitochondria.

Compounds may be screened for mammalian AMPK  $\alpha_1$  antagonist or agonist activity by exposing mammalian AMPK  $\alpha_1$  of the present invention to the compounds and assessing the activity of the mammalian AMPK  $\alpha_1$ . Suitable screening methods for identifying compounds which regulate the activity of mammalian AMPK  $\alpha_1$  include any conventional assay systems for determining such effects. For example, a peptide containing a

serine residue exclusively phosphorylated by AMP protein kinase is incubated in the presence of a preparation of AMP protein kinase and a radiolabel such as gamma  $^{32}\text{P}$ [ATP]. The reaction is allowed to proceed for a period of about 5 minutes and is  
5 conveniently terminated by the addition of acid. The phosphorylated peptide is conveniently separated from unincorporated radiolabel by binding to a charged membrane following washing. The degree of phosphorylation of the peptide is a measure of the activity of the mammalian AMPK  $\alpha_1$ .

10 In addition, compounds may be screened for ability to regulate expression of mammalian AMPK  $\alpha_1$  in a cell by exposing the cell transformed with the polynucleotide of the first aspect of the present invention to the compound and assessing the level of expression of the polynucleotide encoding  
15 mammalian AMPK  $\alpha_1$ . Suitable screening methods for identifying compounds which regulate expression of mammalian AMPK  $\alpha_1$  include those which involve the detection and/or determination of the amount of mammalian AMPK  $\alpha_1$  or messenger RNA that encodes mammalian AMPK  $\alpha_1$  or protein in the presence of the  
20 relevant test compound.

As indicated above, the compounds which regulate activity of mammalian AMPK  $\alpha_1$  are considered to be of potential use in the treatment of, for example, hypercholesterolemia, hyperlipidemia, obesity, clinical syndromes associated with  
25 hypoxia or ischemia (e.g., myocardial infarction, stroke), disorders of nutrition and diabetes mellitus.

In a fifth aspect, the present invention provides an antibody which binds selectively to a polypeptide according to the fourth aspect of this invention. By "binds selectively" it  
30 is meant that the antibody does not bind to mammalian AMPK  $\alpha_2$  as defined in Figure 3A of WO 94/28116. The antibody may be a polyclonal or monoclonal antibody. It will be understood that antibodies of the present invention may be used in a number of procedures. These include monitoring protein expression in  
35 cells; the development of assays to measure kinase activity; and the precipitation of AMP protein kinase and associated proteins which may lead to characterization of these proteins.

Full-length cDNAs for the mammalian AMPK  $\beta$  and AMPK  $\gamma$  subunits have now been cloned. These clones have been used to characterize the tissue distribution of subunit mRNA and to express subunit protein in both bacteria and mammalian cells.

5 Identification of their complete sequences has also led to the identification of protein families for each of these non-catalytic subunits.

Accordingly, in a sixth aspect, the present invention provides an isolated polynucleotide which encodes mammalian

10 AMPK  $\beta$ , the polynucleotide comprising a nucleic acid sequence of SEQ ID NO: 61. Also provided are vectors comprising such a polynucleotide, host cells transformed with such a vector and recombinant proteins encoded by such a polynucleotide.

In a seventh aspect, the present invention provides

15 a method of producing mammalian AMPK  $\beta$  which comprises culturing the cell including the polynucleotide of the sixth aspect of the present invention under conditions which allow expression of the polynucleotide encoding AMPK  $\beta$  and recovering the expressed AMPK  $\beta$ .

20 In an eighth aspect, the present invention provides a substantially purified polypeptide, the polypeptide having an amino acid sequence of SEQ ID NO: 62.

In a ninth aspect, the present invention provides an isolated polynucleotide which encodes mammalian AMPK  $\gamma$ , the

25 polynucleotide comprising a nucleic acid sequence of SEQ ID NO: 63. Also provided are vectors comprising such a polynucleotide, host cells transformed with such a vector and recombinant proteins encoded by such a polynucleotide.

In a tenth aspect, the present invention provides a

30 method of producing mammalian AMPK  $\gamma$  which comprises culturing the cell including the polynucleotide of the ninth aspect of the present invention under conditions which allow expression of the polynucleotide encoding AMPK  $\gamma$  and recovering the expressed AMPK  $\gamma$ .

35 In an eleventh aspect, the present invention provides a substantially purified polypeptide, the polypeptide comprising an amino acid sequence of SEQ ID NO: 64.



### Detailed Description of the Invention

Mammalian AMPK, as isolated from rat and porcine liver, contains three polypeptide subunits, termed AMPK  $\alpha$ , AMPK  $\beta$  and AMPK  $\gamma$ . The  $\alpha$  subunit contains the kinase catalytic domain sequence and is highly homologous to several members of the SNF1 kinase family. Multiple isoforms of the  $\alpha$  subunit have now been identified with  $\alpha_1$  being responsible for about 90% of the AMPK activity detected in liver extracts. In addition, it has now been established that full-length AMPK  $\beta$  and AMPK  $\gamma$  subunits are likewise homologous to two classes of proteins in *S. cerevisiae*. This extends information previously available from limited peptide sequence analysis and from smaller PCR-derived cDNAs (Stapleton et al. (1994) *J. Biol. Chem.* 269, 29343-29346). Further, by cDNA cloning and direct peptide sequencing it has been demonstrated which isoforms of AMPK  $\beta$  and AMPK  $\gamma$  subunits interact with the  $\alpha_1$  catalytic subunit in liver. Thus, it has now been found that these non-catalytic subunits, like the  $\alpha$  subunit isoforms, have a wider tissue distribution, as evidenced by mRNA content of several rat tissues, than expected from the AMPK activity distribution previously reported by Gao et al. (1995) *Biochem. Biophys. Acta.* 1200, 73-82 and Davies et al. (1989) *Eur. J. Biochem.* 186, 123-128.

A novel isoform of the mammalian AMPK catalytic subunit has now been identified and is referred to herein as AMPK  $\alpha_1$ . The  $\alpha_1$  (548 residues) and  $\alpha_2$  (552 residues) isoforms of AMPK have 90% amino acid sequence identity within the catalytic core but only 61% elsewhere. The major differences in the  $\alpha_1$  and  $\alpha_2$  sequences occur in their COOH-terminal tails which suggests that they may interact with different proteins within this region.

It has now been found that the  $\alpha_2$  8.5 kb mRNA is most abundant in skeletal muscle with lower levels in liver, heart and kidney. In contrast, very low levels of the  $\alpha_1$  6 kb mRNA were found in all tissue except testis, where a low level of an uncharacterized 2.4 kb mRNA was observed. The low levels of  $\alpha_1$  mRNA explains why the  $\alpha_1$  isoform was more difficult to clone

than the  $\alpha_2$  isoform. The  $\alpha_1$  isoform of the AMPK catalytic subunit, however, accounts for approximately 94% or more of the SAMS peptide phosphotransferase activity of rat liver and is therefore the predominant active expressed hepatic isoform.

5           A series of synthetic peptides including analogues of proteins not known to be substrates for the AMPK were screened with partially purified enzyme (purified to the DE-52 step). These included the myosin light chains, ADR1, glycogen synthase and phospholemman. The phospholemman peptides tested were poor  
10 substrates and not investigated further. The glycogen synthase peptide, PLSRTL<sup>S</sup>VAAKK (SEQ ID NO: 46) was phosphorylated in an AMP-dependent manner at approximately 40% of the rate of the SAMS peptide, however, this peptide is an excellent substrate for a number of protein kinases, including protein kinase C and  
15 calmodulin dependent protein kinase II (Kemp, B.E. and Pearson, R.B. (1991) in *Protein Phosphorylation*, Hunter, T and Sefton, B.M. (eds) *Methods in Enzymology*, 200, 121-134). The myosin light chain peptides tested were phosphorylated with rates approximately 15% of the SAMS peptide. It was found that the  
20 ADR1 peptides ADR1(225-234) and ADR1(222-234)<sup>P229</sup> were phosphorylated at rates of approximately 50% of the SAMS peptide. Results from these experiments indicate that the ADR1(222-234)<sup>P229</sup> peptide is phosphorylated with an apparent  $K_m$  of approximately 3  $\mu M$  compared to 33  $\mu M$  for the SAMS peptide.

25           In view of the low  $K_m$  of the ADR1(222-234)<sup>P229</sup> peptide as a substrate for the AMPK, affinity purification of the enzyme with this peptide was attempted. Initially the peptide was coupled to CNBr-activated sepharose. Although the peptide linked sepharose bound the AMPK containing fractions the enzyme  
30 could not be differentially eluted from contaminating proteins with salt gradients. In contrast when the ADR1(222-234)<sup>P229</sup> peptide was coupled to Pharmacia HiTrap column the AMPK was bound very tightly and required 2 M NaCl plus 30% ethylene glycol to elute it. Because the enzyme bound so tightly to  
35 this substrate affinity column it was possible to load the enzyme in buffer containing 0.5 M NaCl. The resultant purified AMPK consisted of a 63 kDa catalytic subunit and 40 kDa and 38

kDa subunits related to sip2 and snf4, respectively. In some preparations the AMPK was associated with high molecular weight material that corresponded to non-muscle myosin as assessed by tryptic peptide sequencing. An apparent purification of up to 5 38,000 with a yield of 15% and a recovery of 90  $\mu$ g of enzyme was obtained. The fold purification may be an overestimate due to the presence of uncharacterized inhibitory material in the early fractions. The enzyme was not apparent on SDS-PAGE until the final step of purification. The avidity of the enzyme for 10 the peptide bound to the Pharmacia HiTrap resin was greater than could be expected from the free peptide binding to the enzyme ( $K_m$  3  $\mu$ M). Since the peptide linked to sepharose did not bind the enzyme as tightly it seems reasonable that the enhanced binding is due in part to the aminohexanoic acid 15 linker between the peptide and the resin. In the case of the cAMP-dependent protein kinase there is a hydrophobic pocket between the D and G helices that is responsible for high affinity binding of the peptide inhibitor PKI. Since the ADR1(222-234)P229 peptide, LKKLTLRASFSQA (SEQ ID NO: 47), is 20 linked through the amine residues on its N-terminus or Lys residues, it is possible that the hydrophobic linker group has been fortuitously juxtaposed to a hydrophobic pocket on the AMPK.

In the course of sequencing the porcine AMPK it was 25 found that the amino acid sequence of some peptides derived from the pig liver AMPK  $\alpha$  subunit did not match those deduced from the rat liver cDNA sequence (Carling et al. (1994) *J. Biol. Chem.* 269, 11442-11448; Gao et al. (1995) *Biochem. Biophys. Acta.* 1200, 73-82). Therefore, the rat liver AMPK 30 catalytic subunit,  $\alpha$  was purified and peptides accounting for 40% of the protein sequenced (222/548 residues, SEQ ID NOS: 27-43). Eight of the 16 peptides contained mismatched residues with the reported AMPK cDNA sequence, but did match the pig liver enzyme sequence (SEQ ID NOS: 13-26). Using RT-PCR and 35 cDNA library screening, a cDNA sequence of the rat hypothalamus enzyme was obtained that accounted for all of the peptide sequences of the purified rat liver AMPK catalytic subunit

- 10 -

containing mismatches. The cDNA sequence of this AMPK catalytic subunit has been named  $\alpha_1$ , since it corresponds to the purified enzyme and is clearly derived from a different gene than the previously cloned a sequence (now referred to as  $\alpha_2$ ). The  $\alpha_1$  isoform of the AMPK catalytic subunit accounts for approximately 94% or more of the SAMS peptide phosphotransferase activity of rat liver and is therefore the predominant active expressed hepatic isoform. Despite sequencing multiple preparations of the AMPK catalytic subunit from both pig and rat liver (SEQ ID NOs 13-26 and 27-43, respectively), no peptides were obtained that matched the  $\alpha_2$  isoform sequence.

Within the catalytic cores of the  $\alpha_1$  and  $\alpha_2$  isoforms, there is 90% amino acid identity but only 61% identity outside the catalytic core. Strong homology between the  $\alpha_1$  and  $\alpha_2$  sequences in the vicinity of the substrate binding groove, inferred from the protein kinase crystal structure for positions P<sub>1</sub> to P<sub>5</sub>, suggest that the substrate specificities will be related. The substrate anchoring loop (also called the lip or activation loop) contains an insert FL<sup>170</sup> for  $\alpha_1$ ,  $\alpha_2$  and snflp that may provide a hydrophobic anchor for a P<sub>3</sub> or P<sub>4</sub> hydrophobic residue in the peptide substrate. There is also E<sup>100</sup> (E<sup>127</sup> in cAMP-dependent protein kinase) and D<sup>103</sup> available for a P<sub>3</sub> basic residue specificity determinant for both the  $\alpha_1$ ,  $\alpha_2$  and snflp. Both isoforms contain a Thr-172 residue equivalent to Thr-197 in the cAMP-dependent protein kinase, which is likely to be phosphorylated and necessary for optimal activity. Since the major differences in the  $\alpha_1$  and  $\alpha_2$  sequences occur in their COOH-terminal tails they may interact with different proteins within this region.

Northern blot analysis of the  $\beta$  and  $\gamma$  subunits revealed a complex pattern of expression. The  $\beta$  subunit mRNA was least abundant with similar levels across a range of tissues except brain, whereas the  $\gamma$  subunit mRNA was abundant in heart, lung, skeletal muscle, liver and kidney. An earlier report on the tissue distribution of the AMPK activity had claimed that it was predominantly a liver enzyme (Davies et al.

(1989) *Eur. J. Biochem.* 186, 123-128). In view of the mRNA distribution of the  $\alpha_1$  and  $\beta$  subunits, the tissue distribution of the AMPK activity was reassessed. The kidney contained the highest specific activity with similar levels in the liver, lung and heart and little, if any, activity in skeletal muscle. It is clear that the AMPK activity has a wider tissue distribution than appreciated heretofore, closely paralleling the distribution of  $\alpha_1$  mRNA and not that of  $\alpha_2$  mRNA. Using peptide specific antisera to  $\alpha_1$  (residues 339-358) and  $\alpha_2$  (residues 352-366) it was found that the  $\alpha_2$  immunoreactivity was predominant in the heart, liver and skeletal muscle where there is also the highest concentrations of  $\alpha_2$  mRNA. In contrast the  $\alpha_1$  immunoreactivity is widely distributed as is the less abundant  $\alpha_1$  mRNA. The antibody to  $\alpha_2$  recognized a minor component in the purified  $\alpha_1$  preparation but sufficient amounts of this have not been obtained to determine whether it represents weak cross reactivity with a form of  $\alpha_1$ , an additional isoform of the AMPK or a low level contaminant of the  $\alpha_1$  preparation by the  $\alpha_2$  isoform. The antibody to  $\alpha_2$  does not immunoprecipitate  $\alpha_1$  activity from affinity purified  $\alpha_1$  AMPK. Both  $\alpha_1$  and  $\alpha_2$  migrate on SDS-PAGE at approximately 63 kDa. It was also found that the liver  $\alpha_2$  immunoreactivity was not bound by the peptide substrate affinity column. This column specifically binds the  $\alpha_1$  isoform. Using immune precipitation of the effluent from the peptide substrate affinity column with  $\alpha_2$  specific antibody it was found that the  $\alpha_2$  isoenzyme contained  $\beta$  and  $\gamma$  subunits and catalyzed the phosphorylation of the SAMS peptide. Immune precipitates of  $\alpha_1$  and  $\alpha_2$  showed variable activation by 5'-AMP ranging from 2-3 and 3-4 fold, respectively. There was also an approximate 60 kDa band recognized by the  $\alpha_1$ -specific antibody in tissue extracts from heart and lung. This band is not present in the purified liver enzyme and its relationship to the  $\alpha_1$  isoform is not yet known.

The proportion of SAMS peptide phosphotransferase activity bound to the peptide affinity column with a single pass varied (ranged 90-92%, n=7 and 74-86%, n=6 rat liver

preparations). With recycling, approximately 94% of the activity was bound to the column. The residual activity was attributable to  $\alpha_2$  isoform activity based on immunoprecipitation with the  $\alpha_2$ -specific antibody. However, 5 the amount of protein immunoprecipitated based on Coomassie blue staining indicated that there was substantially more  $\alpha_2$  protein than was expected from only 6% of the total SAMS peptide activity. The apparent specific activity of the isolated rat hepatic AMPK  $\alpha_2$  isoform with either the SAMS 10 peptide or acetyl CoA carboxylase as substrate was more than 20-fold lower than the AMPK  $\alpha_1$  isoform. This estimate is based on measurements using the  $\alpha_2$  enriched fraction ( $\alpha_1$  depleted) and quantitation by immunoblotting compared to bacterially expressed  $\alpha_2$ .

15 The specific activity of the purified  $\alpha_2$  isoform is not yet known in the absence of bound antibody. Based on the  $\alpha_2$  cDNA sequence, Carling et al. (1994) *J. Biol. Chem.* 269, 11442-11448 reported that a peptide specific antibody immunoprecipitated virtually all of the partially purified AMPK 20 activity from liver. The peptide used in their experiments, PGLKPHPERMPPLI (SEQ ID NO: 48), contains 8/15 residues that are identical (underlined) between  $\alpha_1$  and  $\alpha_2$ , so it seems reasonable that their polyclonal antisera may recognize both isoforms.

These experiments make clear that there is an 25 isoenzyme family of AMPK  $\alpha$  catalytic subunits, thus increasing the complexity of activity analysis. This also raises the question of what function the  $\alpha_2$  isoform has and whether  $\alpha_2$  associates with a specific subset of  $\beta$  and  $\gamma$  subunits. A significant fraction of the  $\alpha_2$  isoform mRNA has a 142 bp out- 30 of-frame deletion within its catalytic domain that would encode a truncated, non-functional protein (Gao et al. (1995) *Biochem. Biophys. Acta.* 1200, 73-82; Verhoeven et al. (1995) *Eur. J. Biochem.* 228, 236-243). The close sequence relationship between the  $\alpha_1$  isoforms from pig, rat and human means that 35 there is strong conservation across species. Previously, it was reported that human liver does not contain AMPK mRNA (Aguan et al. (1994) *Gene* 149, 345-350). However, it is now clear that

$\alpha_2$  mRNA was being probed for and not the dominant  $\alpha_1$  isoform mRNA. The gene encoding the human liver AMPK catalytic subunit reported on chromosome 1 is therefore the gene for the  $\alpha_2$  isoform whereas the gene for the  $\alpha_1$  isoform is located on  
5 chromosome 5. The AMPK subunit genes have now been mapped predominantly to the following chromosomal locations:  $\alpha_1$ , 5p12;  $\beta$ , 5q24.1; and  $\gamma$ , 12q13.1.

Recent genome sequencing has revealed multiple isoforms of the non-catalytic  $\gamma$  and  $\beta$  subunits of the AMPK.  
10 There appear to be at least three isoforms of the  $\gamma$  subunit in brain with  $\gamma_2$  and  $\gamma_3$  present, distinct from the rat liver  $\gamma_1$  isoform. Human brain also contains multiple  $\beta$  subunit isoforms distinct from the rat liver  $\beta_1$  isoform. The accession numbers for putative AMPK  $\beta$  and  $\gamma$  subunit isoforms are  $\gamma_2$ ,  
15 M78939;  $\gamma_3$ , R52308;  $\beta_2$ , R20494 and  $\beta_3$ , R14746. Thus, a potentially large subfamily of heterotrimeric AMPKs, based on various combinations of all three AMPK subunits, may be present.

The structural relationships between the AMPK and SNF1  
20 kinase, as well as the presence of multiple isoforms, brings into focus a vista of questions concerning the diverse physiological roles of this new subfamily of protein kinases. Whereas the AMPK regulates lipid metabolism in hepatocytes under conditions of metabolic stress, its role in other  
25 tissues, including the heart and kidney, are unknown. Recent studies have shown that the AMPK is activated during cardiac ischaemia, and the activation persists during reperfusion, possibly contributing to the ischaemia-driven decoupling of metabolism and cardiac mechanical function (Kudo et al. (1995)  
30 *J. Biol. Chem.* 270, 17513-17520).

Regulation of cardiac acetyl-CoA carboxylase by AMPK plays an important role in the switching of cardiac metabolism between the use of glucose and fatty acids as oxidative fuel. In the  $\beta$  cell of the pancreas, where AMPK subunits are highly  
35 expressed in islet cells, glucose availability rapidly regulates acetyl-CoA carboxylase through changes in AMPK-directed phosphorylation, suggesting strongly a role for AMPK

in stimulus-secretion coupling for insulin release. In addition to these metabolic roles, members of the SNF1 protein kinase subfamily appear to play important roles in development, with the par-1 gene of *C. elegans* playing an essential role in  
5 embryogenesis.

PCR amplification of pig and rat liver cDNA with degenerate oligonucleotides corresponding to selected AMPK  $\beta$  peptide sequences yielded two major PCR products (Stapleton et al. (1994) *J. Biol. Chem.* 269, 29343-29346). One product, a  
10 rat 309 bp partial length cDNA, was used to screen a rat liver cDNA library, yielding a 1107 bp clone (SEQ ID NO: 61). The screening PCR probe corresponded to nt residues 279-588 of this sequence. This clone contains an open reading frame encoding for a 270 amino acid peptide (SEQ ID NO: 62), which contains  
15 all of the 15 independent (some overlapping) peptide sequences obtained from extensive sequence analysis of the purified protein. The translational start methionine codon is assigned from the typical Kozak sequence present for a initiation codon and the lack of any other upstream in-frame methionine codons.  
20 While no in-frame stop codon is present in this 5'-upstream sequence, a human expressed sequence tag (EST) cDNA (GenBank accession no. T78033) in the database contains such a stop codon preceding the same assigned methionine start. This reading frame, however, predicts a protein of 30,464 daltons,  
25 well below the estimated molecular weight of 40 kDa evident on SDS gel electrophoresis.

In order to clarify the size of the protein product that could be synthesized from this cDNA, the AMPK  $\beta$  clone was expressed both in bacteria and mammalian cells. In both  
30 expression systems, the protein product migrates at a higher than predicted molecular weight. When purified as a His<sup>6</sup>-tagged fusion protein from *E. coli*, the isolated protein migrates on SDS gels with an apparent molecular weight of about 43,000 Da (the same as the ovalbumin standard). This  
35 corresponds to a AMPK  $\beta$  polypeptide product of 40 kDa with an additional 3 kDa daltons of fusion tag sequence derived from the pET vector. When expressed in mammalian cells from an HA-



tagged expression vector, two polypeptides are evident with the major product corresponding to a 40 kDa species (after correction for the size of the HA epitope tag). A second product of 42-43 kDa is also evident using this expression  
5 system. Taken together, these data demonstrate that the protein product of this AMPK  $\beta$  migrates on SDS-PAGE with an anomalously high molecular weight.

Comparison of the rat liver AMPK  $\beta$  sequence to the database reveals that it is highly homologous to three yeast  
10 proteins (Sip1p, Sip2p and Gal83p) and to two recently cloned human EST-cDNA sequences. This alignment, as gapped according to the sequence of the *S. cerevisiae* protein, Sip1p (Yang et al. (1992) *Science*, 257, 680-682), is most striking at the C-terminus of AMPK  $\beta$  and these yeast proteins.

15 The AMPK  $\beta$  subunit is a mammalian homolog of a class of proteins in yeast, represented by Sip1p/Sip2p/Gal83p. The GAL83 gene product is known to affect glucose repression of the GAL genes. All of these proteins have been shown to interact with the Snf1p protein kinase either in the 2-hybrid system or  
20 by immunoprecipitation. It has been proposed that these proteins serve as adaptors that promote the activity of Snf1p toward specific targets. Based on analysis of yeast mutants, it has been suggested that these proteins may facilitate interaction of Snf1p with unique and different targets. Of  
25 interest is the demonstration of a highly conserved domain of about 80 amino acids in the C-terminus of Sip1p/Sip2p/Gal83p, termed the ASC domain (association with Snf1p complex) (Yang et al. (1994) *EMBO J.* 13, 5878-5886). As studied in the 2-hybrid system, the ASC domain of both Sip1p and Sip2p interacts  
30 strongly with Snf1p. However, the interaction of Sip2p with Snf1p is not entirely lost on deletion of this domain, suggesting that the ASC domain is not solely responsible for this protein-protein interaction. A putative ASC domain is also highly conserved in the C-terminus of rat liver AMPK  $\beta$  (aa  
35 residues 203-270), suggesting that this region may be responsible, in part, for binding to the AMPK  $\alpha$  subunit.

- 16 -

AMPK  $\beta$ , like Sip2p and Gal83p, is phosphorylated in vitro when associated with a catalytic subunit (AMPK  $\alpha$  or Snflp, respectively). Mutations of Gal83p can abolish most of the Snflp kinase activity detectable in immune complexes, 5 precipitated with anti-Snflp antibody. A *Sip2p/E gal 83/E* mutant shows reduced Snfl protein kinase activity, that is restored following expression of either Sip2p or Gal83p LexA-fusion proteins in the mutant strain (Yang et al. (1994) *EMBO J.* 13, 5878-5886). Taken together, these data suggest the 10 possibility that AMPK  $\beta$  may also serve as an adaptor molecule for the AMPK  $\alpha$  catalytic unit and will positively regulate AMPK activity.

AMPK  $\beta$  appears to have anomalous migration on SDS gels, with the polypeptide migrating at a  $M_r$  approximately 10 15 kDa larger than the size predicted from the cDNA. This slower migration is evident for both the bacterially expressed His<sup>6</sup>-fusion protein and for the protein expressed in COS7 cells. These observations suggest that higher orders of structure are responsible for the anomalous migration on SDS-PAGE. The AMPK 20  $\beta$  subunit is autophosphorylated in vitro; this suggests that the two AMPK  $\beta$  bands expressed on transfection of mammalian cells with AMPK  $\beta$  cDNA may result from a similar post-translational modification giving rise to smaller mobility shifts. Interestingly, this aberrant migratory behavior of 25 AMPK  $\beta$  is similar to that of its yeast homolog, Gal83p. The LexA-fusion protein(s) of Gal83, as expressed in yeast, also migrate at greater than the expected molecular weight and display more than one band on SDS gels, consistent with the known phosphorylation of Gal83p by Snflp. Mass spectrometry 30 analysis of the  $\beta$ -subunit indicates that the amino terminal glycine is myristylated and that the subunit is isolated in mono- and di- phosphorylated forms.

Using the MOPAC procedure and other PCR amplification protocols, a 192 bp cDNA corresponding to rat liver AMPK  $\gamma$  35 sequence was obtained and used for library screening to obtain a partial length rat liver cDNA of approximately 1.3 kb. This sequence did not contain either a start methionine codon or all

the peptide sequences obtained from the purified protein. Attempts to extend this sequence to the 5'-end by the use of a primer extension library and 5'-RACE only succeeded in adding about 200 nt to this sequence without identification of the  
5 start codon. A partial length rat cDNA was then used to screen a human fetal liver library, which did yield the full-length clone depicted in SEQ ID NO: 63. This clone contains a deduced amino acid sequence (SEQ ID NO: 64) corresponding to all of 22 independent (some overlapping) peptide sequences obtained from  
10 the purified rat and porcine liver AMPK  $\gamma$ , confirming clonal identity.

A typical Kozak translation initiation sequence surrounds the assigned methionine start codon; this start is also in-frame with a 5'-upstream stop codon. The assigned  
15 start methionine is followed by an open reading frame predicting a protein of 331 amino acids and of 37,546 Da, which corresponds to the molecular weight observed on SDS gel electrophoresis of the protein as purified from rat and porcine liver. Expression of a truncated rat AMPK  $\gamma$  cDNA (aa residues  
20 33-331) and the full-length human AMPK  $\gamma$  (331 aa) in COS7 cells yields products consistent with the molecular weight predicted for each cDNA (34,081 and 37,577 daltons, respectively). The rat liver AMPK  $\gamma$  product expressed in bacteria also displayed the molecular weight predicted by the cDNA. Thus, unlike AMPK  
25  $\beta$ , there is no anomalous migration of the protein product of AMPK  $\gamma$  cDNA.

Comparison of the human and rat liver AMPK  $\gamma$  amino acid sequences to the database yields a significant alignment of this protein with the *S. cerevisiae* Snf4p. In addition,  
30 human full-length cDNA of the present invention also aligns with several other human partial length EST-cDNA sequences from brain, breast, placenta, liver and heart, recently reported in the database. Inspection of these sequences reveals that there are multiple isoforms of the human AMPK  $\gamma$   
35 protein. There are likely also similar AMPK  $\gamma$  isoform families expressed in the rat and pig. This latter expectation is based on sequence analysis of 14 other MOPAC-derived partial AMPK  $\gamma$

cDNA sequences, as identified on colony hybridization of the AMPK  $\gamma$  MOPAC products with  $^{32}$ P-labeled degenerate oligonucleotides. These products showed at least two reproducible patterns of nucleotide heterogeneity within the non-degenerate core.

Rat and human liver AMPK  $\gamma$  is a mammalian homolog of the *S. cerevisiae* Snf4p (CAT3) (Celenza et al. (1989) *Mol. Cell. Biol.*, 9, 5045-5054; Schuller, H.J. and Entian, K.D. (1988) *Gene*, 67, 247-257; Fields, S. and Song, O.K. (1989) *Nature*, 340, 245-246). Snf4p was shown to interact with the Snf1p protein in the first reported use of the 2-hybrid system and also co-immunoprecipitates with it (Haygood, M.G. (1993) *Biotechniques* 15, 1084-1089). Indeed, on isolation of the Snf1p kinase from yeast, Snf4p, but not the other Snf1p-interacting proteins, co-purifies in a 1:1 stoichiometry with the Snf1p polypeptide. Analysis of SNF4 mutants in yeast suggests that Snf4p also positively regulates the activity of its associated catalytic subunit, Snf1p. By analogy, our prediction is that AMPK  $\gamma$  will also have such a positive influence on the AMPK  $\alpha$  subunit.

Examination of the database reveals that, in addition to the homology of AMPK  $\gamma$  to Snf4p, there are 2 or 3 different human proteins highly homologous or identical to our human and rat liver AMPK  $\gamma$  sequences. However, some of these database sequences, as predicted from EST-cDNAs in brain, heart, breast and placenta, are distinct from each other and from our clones; some, for example, have a C-terminal extension. This indicates that there is a mammalian isoform family of potential AMPK  $\gamma$  subunits, each perhaps with different tissue expression and regulatory roles. It is suggested that these different gamma isoforms be designated  $\gamma_1$ ,  $\gamma_2$ ,  $\gamma_3$ , ...,  $\gamma_n$ , as their full-length sequences are delineated. The rat liver/human liver AMPK  $\gamma$  sequence of the present invention is designated herein as AMPK  $\gamma_1$ .

AMPK  $\alpha$  catalytic unit is widely expressed in several rat tissues. AMPK  $\beta$  and AMPK  $\gamma$  sequences have a similar wide tissue expression. Two species of AMPK  $\gamma$  mRNA of 2.7 and 1.9

kb are evident in total mRNA preparations; only the latter is present in polyA<sup>+</sup>-RNA from rat liver, suggesting that the larger mRNA is an unprocessed precursor. Only a single mRNA species for AMPK  $\beta$  of 1.9 kb is evident. Both AMPK  $\gamma$  and AMPK  $\beta$  mRNAs are highly expressed in kidney, white adipose tissue, lung and spleen, while AMPK  $\gamma$  mRNA appears to be more highly expressed in heart and brain. While detectable, the mRNA level for each subunit is relatively lower in skeletal muscle, lactating mammary gland and liver. In other studies, high concentrations of mRNA have been found for both subunits in the rat Fao hepatoma cell and the Syrian hamster insulin-secreting HIT cell, cell lines that both express substantial levels of AMPK activity.

AMPK was first recognized as a protein kinase active on enzymes of lipid metabolism (acetyl-CoA carboxylase, HMG Co-A reductase and hormone-sensitive lipase). However, as has been observed for the AMPK  $\alpha$  subunit, the AMPK  $\beta$  and AMPK  $\gamma$  subunits have wider tissue distribution than might be expected for a protein active only in the regulation of lipid metabolism. While mRNAs for each are detectable in "classic" lipogenic tissues like liver, white adipose tissue and lactating mammary gland, high concentrations of mRNA in non-lipogenic tissues like heart, brain, spleen and lung, for example, suggest that these proteins have roles that extend beyond the regulation of biosynthesis of fatty acids and sterols and fatty acid oxidation.

For example, the striking homology of all three subunits to yeast proteins that are involved in nutrient (glucose) responses suggests that the three mammalian proteins may be involved in glucose (or other nutrient) regulation of gene expression in mammalian tissues or in other adaptive responses to a changing nutrient environment. In addition, AMPK may be an important "metabolic sensor" linked to oxidative fuel choice in the heart and to glucose sensing in the pancreatic beta cell, perhaps being important for insulin secretion.

The following nonlimiting examples are provided to further illustrate the instant invention.

#### EXAMPLES

##### Example 1: Purification of AMPK Catalytic Subunit ( $\alpha$ 1)

###### 5 Enzyme Purification

AMPK was purified from porcine liver. Liver (1 kg) was homogenized in 4,000 ml of buffer. A 2.5-7.0% (w/v) PEG 6000 fraction was prepared and the resultant fraction batched onto 1,500 ml of DEAE cellulose (Whatman, Clifton, NJ) and  
10 eluted with buffer containing 0.25 M NaCl (2,000 ml). The eluate was chromatographed on 150 ml Blue Sepharose (Pharmacia, Uppsala, Sweden) and the AMPK eluted with buffer containing 1 M NaCl. The enzyme fraction was concentrated and desalted by  
15 10% (w/v) PEG-6000 precipitation prior to chromatography by peptide substrate affinity chromatography. The peptide substrate affinity column was washed with the same buffer containing 0.1% (v/v) Triton X-100 and 0.5 M NaCl and the AMPK eluted with this buffer containing 2 M NaCl and 30% (v/v) ethylene glycol.

###### 20 Protein kinase assays

The AMPK was assayed in accordance with procedures described by Davies et al. (1989) *Eur. J. Biochem.* 186, 123-128 using the SAMS peptide substrate, HMRSAMSGHLVKRR-amide (SEQ ID NO: 49). The enzyme was diluted in diluting buffer (20 mM  
25 HEPES pH 7.0, 0.1% (v/v) Triton X-100) prior to assay and the reactions were initiated by adding 10 ml diluted enzyme to the reaction mixture containing peptide substrate. The reactions were stopped by withdrawing 30 ml aliquots and applying to P81 papers in accordance with procedures described by Pearson,  
30 R.E., Mitchelhill, K.I., and Kemp, B.E. (1993) in *Protein Phosphorylation: A Practical Approach*, Hardie, G.D. (ed) Oxford University Press, pp 265-291.

### Peptide synthesis

Peptides were synthesized using an Applied Biosystems 430 synthesizer in accordance with procedures described by Pearson, R.B., Mitchelhill, K.I., and Kemp, B.E. (1993) in *Protein Phosphorylation: A Practical Approach*, Hardie, G.D. (ed) Oxford University Press, pp 265-291. All peptides were purified by cation-exchange chromatography followed by reverse phase chromatography. Peptides were analyzed by quantitative amino acid analysis using a Beckman 6300 amino acid analyzer. The peptide substrate affinity column was prepared by coupling the ADR1(222-234)<sup>F229</sup> peptide to a Pharmacia HiTrap N-hydroxysuccinamide ester activated superose column. This resin contains a 6-aminohexanoic acid spacer arm. The conditions of coupling were performed in accordance with manufacturer's instructions with 10 mg peptide per 5 ml column and peptide coupling was monitored by reverse phase HPLC.

### Example 2: Isolation of cDNA Encoding AMPK Catalytic Subunit ( $\alpha$ 1)

#### Peptide Sequencing

Peptides were derived from rat and porcine  $\alpha$ 1 subunit of the AMPK, by *in situ* proteolysis in accordance with procedures described by Mitchelhill et al. (1994) *J. Biol. Chem.* 269, 2361-2364 and sequenced on either an Applied Biosystems 471A Protein Sequencer or a Hewlett Packard G1000A Protein Sequencer.

#### Tissue Distribution Activity Studies

A 35% saturated ammonium sulfate fraction was prepared for each tissue, following homogenization in AMPK homogenization buffer (HB, 50 mM Tris-HCl pH 8.5, 250 mM sucrose, 5 mM sodium pyrophosphate, 50 mM sodium fluoride, 1 mM EGTA, 1 mM EDTA, 1 mM DTT, 1 mM benzamidine, 1  $\mu$ g/ml soybean trypsin inhibitor and 0.2 mM phenylmethyl-sulfonylfluoride). The resultant pellet was resuspended in 5 ml HB and assayed for protein concentration. The AMPK was assayed in accordance with procedures described by Mitchelhill et al. (1994) *J. Biol.*

Chem. 269, 2361-2364 with the following modifications: a final reaction volume of 120  $\mu$ l was used, enzyme aliquots (30  $\mu$ l) containing 1  $\mu$ g protein pre-diluted in 50 mM Tris-HCl pH 7.5 and 0.05% (v/v) Triton X-100 were used to initiate the reaction. Three aliquots (30  $\mu$ l) were removed after 2, 4 and 6 min. Reactions were performed in duplicate  $\pm$  5'-AMP (200  $\mu$ M), with a minus peptide substrate control. The specific activity of the enzyme was determined using linear rates of phosphorylation with the specific synthetic peptide substrate SAMS. The AMPK was purified from rat or porcine liver as described in Example 1 using substrate affinity chromatography.

#### *Isolation of AMPK cDNA*

A radiolabelled cDNA (774bp) encoding porcine AMPK  $\alpha_1$  was used to screen a rat hypothalamus Zap II cDNA library (Stratagene, La Jolla, CA) according to the manufacturer's instructions. Positives were plaque-purified on subsequent rounds of screening and phagemid from positive clones were rescued with helper phage (Stratagene). Screening of  $7 \times 10^6$  plaques yielded three unique clones, the largest consisting of an open reading frame, corresponding to AMPK  $\alpha_1$  (2-549).

The AMPK  $\alpha_1$  5' end was isolated using a Gibco 5'-RACE kit (Life Technologies, Grand Island, USA) with an  $\alpha_1$  specific primer to residues 41-48 and rat liver cDNA. Human AMPK  $\alpha_1$  (14-270) was isolated from fetal human liver cDNA primed with sense and anti-sense partially degenerate oligonucleotides to  $\alpha_1$  peptide sequence by RT-PCR. Human AMPK  $\alpha_1$ , residues 291-448 is a partial length human liver cDNA clone obtained from the Lawrence Livermore National Laboratory (clone 78297, accession number T50799).

#### *Northern Blotting*

A rat multiple tissue Northern (MTN) blot (Clontech, Palo Alto, CA, USA) containing 2 mg of poly(A)<sup>+</sup> RNA of individual tissues was probed with <sup>32</sup>P-labelled rat AMPK  $\alpha_1$  and  $\alpha_2$  cDNAs according to the instructions supplied.



### *Production of Anti-AMPK Antibodies*

Polyclonal antibodies to AMPK  $\alpha_1$  and  $\alpha_2$  were prepared as follows. Peptides based on the predicted amino acid sequences of AMPK  $\alpha_1$  for residues 339-358 (DFYLATSPSPDSFLDDHHLTR  
5 (SEQ ID NO: 50)) and AMPK  $\alpha_2$  for residues 352-366 (MDDSAMHIPPGLKPH (SEQ ID NO: 51)) were synthesized and coupled to keyhole limpet hemocyanin (Sigma Chemical Co. St. Louis, MO, H-2133) via a cysteine residue added to the N-terminus of the peptide using the heterobifunctional reagent, N-succinimidyl-3-  
10 (2-pyridyldithio)propionate (Pharmacia, Uppsala, Sweden). New Zealand White rabbits were immunized with 2 mg peptide conjugate initially in 50% (v/v) Freund's complete adjuvant and in 50% (v/v) Freund's incomplete adjuvant for subsequent immunizations. Rabbits were boosted fortnightly with 2 mg  
15 peptide conjugate and bled 7 days after booster injections. Anti-AMPK  $\alpha_1$  and  $\alpha_2$  peptide antibodies were purified by peptide affinity chromatography.

### *Western Blotting*

Multiple rat tissue western blots were prepared as follows. Rat tissues were homogenized in AMPK HB and a 2.5 -  
20 7% polyethylene glycol 6000 fraction was prepared. The resultant pellet was resuspended in 5 ml HB and assayed for protein concentration. One hundred micrograms of each tissue fraction was analyzed by SDS PAGE (13% acrylamide gels);  
25 transferred to nitrocellulose (Schleicher & Schuell, Dassel, Germany); and probed with 3  $\mu$ g/ml and 6  $\mu$ g/ml affinity purified AMPK  $\alpha_1$  and  $\alpha_2$  antibodies, respectively. Primary antibody was detected using anti-rabbit IgG antibody conjugated to horseradish peroxidase (DAKO, Carpinteria, CA, USA) and 0.032%  
30 3,3' -diamino-benzidine (D-5637, Sigma) together with 0.064%  $H_2O_2$ .

### *Purification of AMPK $\alpha_2$*

Affinity purified AMPK  $\alpha_2$  antibody (2 mg) was coupled to CNBr-activated Sepharose 4B (Pharmacia, Uppsala, Sweden)  
35 according to the manufacturer's instructions. The unbound

fraction from the substrate affinity column was applied directly to the AMPK  $\alpha_2$  antibody column, washed with 5 volumes of PBS and eluted with 200 mM glycine buffer pH 2.5 and immediately neutralized.

5 **EXAMPLE 3: Isolation of cDNAs Encoding AMPK Non-Catalytic Subunits**

*AMPK isolation and peptide sequencing*

Porcine and rat liver AMPK was isolated. Peptide sequences derived from the rat liver beta (40 kDa) and gamma  
10 (38 kDa) subunits were obtained after subunit separation by SDS gel electrophoresis, band elution and *in situ* protease digestion in accordance with procedures described by Mitchelhill et al. (1994) *J. Biol. Chem.* 269, 2361-2364 and Stapleton et al. (1994) *J. Biol. Chem.* 269, 29343-29346.

15 *AMPK  $\beta$  subunit cDNA isolation*

Peptide sequences derived from the AMPK  $\beta$  subunit were used to generate partial length AMPK  $\beta$  subunit cDNAs by the polymerase chain reaction (PCR) in accordance with procedures described by Gao et al. (1995) *Biochem. Biophys. Acta.* 1200,  
20 73-82. One product, a 309 bp cDNA, was used to screen a rat liver  $\lambda$ ZAPII cDNA library (Stratagene). Filters were hybridized with  $^{32}$ P-cDNA, labelled with alpha- $^{32}$ P-CTP (3000mCi/mmol, New England Nuclear) by random priming (Random Primer cDNA Labeling System, Gibco/BRL), in 50% formamide, 10X  
25 Denhardt's, 1M NaCl, 50 mM Tris-Cl (pH 7.5), and 100  $\mu$ g/ml salmon sperm DNA at 42°C for 18 hours. They were then washed at room temperature 3 x 10 minutes and then at 55°C for 15 minutes. Autoradiography was for overnight at -80°C. All plates were lifted in duplicate and positive plaques were  
30 purified through 3 additional rounds of plating and re-screening.

*AMPK  $\gamma$  subunit cDNA isolation*

Where peptide sequences are listed herein, the letters Y, H, N and R indicate regions of degeneracy. For the AMPK  $\gamma$  subunit, a 67 bp cDNA was generated by the MOPAC technique described by Lee, C.C. and Caskey, C.T., (1990) in *PCR Protocols*, (Innis, M.A. Gelfand, D.H., Srinisky, J.J., and White, T.J. editors), pp. 46-53, Academic Press, Inc., London. Degenerate PCR primers were synthesized corresponding to the N- and C- terminal sequences of a 17-amino acid rat liver AMPK  $\gamma$  peptide (VVDIYSKFDVINLAAEK (SEQ ID NO: 52)). The sequence of the sense primer was GCGGATCCGTNGAYATHTA (SEQ ID NO: 53) and the sequence of the antisense primer was CGGAATTCYTTYTCNGCNGCNA (SEQ ID NO: 54). BamHI and EcoRI sites were added to the 5'-ends of these primers. The strategy was to create a non-degenerate nucleotide sequence corresponding to the middle portion of the peptide sequence that would be used in library screening. Total rat liver cDNA, prepared with oligo-dT and random hexamers (GIBCO/BRL pre-amplification kit), was used with PCR to amplify a 67-mer (including primers) oligonucleotide corresponding to a portion of the AMPK  $\gamma$  cDNA. The purified PCR product was digested with BamHI and EcoRI and ligated into pBluescript plasmid for transformation of DH5 $\alpha$  bacteria. Colony hybridization was employed to identify clones of interest; colonies were lifted from replica plates onto nitrocellulose filters. Following bacterial lysis and DNA denaturation, filters were probed with a mixture of two <sup>32</sup>P-end-labeled degenerate oligonucleotide probes corresponding to amino acid sequence (KFDVINLA (SEQ ID NO: 55)) internal to that of the two PCR primers. These oligonucleotides (#1: AARTTYGAYGTNATHAAYCTNGC (SEQ ID NO: 56); #2: AARTTYGAYGTNATHAAYTTRGC (SEQ ID NO: 57)) were added in a ratio of two parts oligo #1 to one part oligo #2 to reflect the degeneracy of the Leu codon. Positive colonies were identified and plasmid DNA isolated from each for sequence analysis. One such cDNA was chosen and the non-degenerate "core" 23-mer oligonucleotide sequence was then synthesized for use in library screening (CTCCAAGTTTGATGTTATCAACC (SEQ ID NO: 58)).

- 26 -

Screening of approximately  $10^6$  plaques with this probe, however, did not yield any positive clones.

The non-degenerate 23-mer cDNA was then used in conjugation with degenerate primers constructed from two other peptide sequences to generate a larger AMPK  $\gamma$  cDNA by PCR. Both sense and antisense degenerate oligonucleotide primers corresponding to the peptide sequences, EELQIG (SEQ ID NO: 59) and FPKPEFM (SEQ ID NO: 60), were used together with the sense MOPAC-derived non-degenerate sequence to generate all possible PCR products, using rat liver cDNA as template. The largest product (192 bp) obtained was subcloned in pCR-Script (Stratagene) and sequenced. This sequence, which actually had a predicted amino acid sequence corresponding to all three AMPK  $\gamma$  peptides used in the PCR strategy, was then used for library screening, as above. Screening of  $2 \times 10^6$  plaques with this larger PCR product yielded several positive clones; however, none of the rat cDNAs (1-1.3 kb) isolated corresponded to a full-length open reading frame. In an effort to extend the sequence to the 5'-end of the ORF, a primer extension library was constructed using a AMPK  $\gamma$ -specific antisense primer (Stratagene;  $\lambda$ ZAPII). Additional screening of this library, while yielding some 5'-extended sequence, did not yield the start Met codon. The application of a 5'-RACE strategy with rat liver cDNA was also unsuccessful in attempts at sequence extension, although a 5'-RACE product from porcine liver was obtained. The most 5' rat cDNA sequence (520 bp) was then used to screen a human fetal liver library, which yielded a full-length AMPK  $\gamma$  cDNA.

#### *Plasmid Preparation and DNA sequencing*

Plasmid DNA was prepared using Qiagen Mini- or Midi-columns (Chatsworth, CA) according to the manufacturer's instructions. DNA was sequenced, with vector or gene-specific primers, using an Applied Biosystems Prism(tm) (Foster City, CA) ready reaction Dye Deoxy Terminator Cycle Sequencing kit, and cycled in a Perkin-Elmer PCR Thermocycler, according to the manufacturers' instructions. Dye terminators were removed from

the resulting sequence reactions using a Centri-Step column (Princeton Separations, Inc.). The purified sequencing reactions were then dried in a Speed-Vac and analyzed on an automated DNA sequencer (Applied Biosystems Model 373).

#### 5 Bacterial Expression of cDNAs

Full-length rat AMPK  $\beta$  subunit cDNA and a partial length rat AMPK  $\gamma$  (aa 33-331) subunit cDNA were expressed in *E. coli* using the pET vector system, which introduces polyhistidine (His6) and T7 fusion epitope tag sequences (Novagen, Madison, WI). Bacterial expression was induced with 1.0 mM IPTG at 37°C for 2 hours. Expressed protein was detected by both Coomassie blue staining and immunoblotting with anti-T7 monoclonal antibody (Novagen). The fusion proteins were purified from the inclusion bodies of bacteria by nickel affinity chromatography under denaturing conditions. His6-AMPK  $\beta$  or His6-AMPK  $\gamma$  were solubilized from the inclusion bodies in 6 M urea, according to manufacturer's instructions. After sample application, the column was washed extensively with Tris-Cl (20 mM; pH 7.9), 0.5 M NaCl (0.5 M), imidazole (20 mM) and urea (6 M). The His6-protein was eluted with the same buffer containing 300 mM imidazole.

#### Cellular expression of cDNAs

Full-length rat AMPK  $\beta$  cDNA, a partial length rat AMPK  $\gamma$  (aa 33-331) and full-length human AMPK  $\gamma$  subunit cDNAs were also expressed in COS7 cells. cDNAs were cloned into a pMT2 vector in-frame with a hemagglutinin (HA) epitope tag (pMT2-HA). Transfection was done using Lipofectamine reagent (Gibco/BRL), according to the manufacturer's general protocol. Cells were plated at  $3 \times 10^5$ /well in 6 well plates in DMEM containing 10% fetal calf serum and penicillin/streptomycin. The following day, the cells were switched to serum-free, antibiotic-free DMEM and then lipofectamine-DNA conjugates (2  $\mu$ g of DNA; 10  $\mu$ l lipofectamine per well) diluted in the same medium were added. After 5 hours incubation at 37°C, an equal volume of medium containing 20% fetal calf serum was added to each well. The following morning, the medium was switched to

the original cell medium. Cells were harvested 48 hours after transfection. After washing with PBS, cells were lysed in a buffer containing Tris-Cl (50 mM; pH 7.5), NaCl (100 mM), NaF (50 mM), NaPP<sub>i</sub> (5 mM), EDTA (1 mM), DTT (2 mM) and NP-40 (0.5%) with several protease inhibitors.

For complete lysis, cells were placed on ice for 15 minutes followed by scraping and vigorous vortexing (15 seconds) of the lysate. After clearing of debris by brief centrifugation, this lysate was used for SDS gel electrophoresis and immunoblotting. Blots were probed with an anti-HA monoclonal antibody (derived from the 12CA5 hybridoma line). After secondary probing with an anti-mouse IgG-peroxidase antibody, blots were developed by ECL (Amersham).

#### *Northern Blot Analysis*

Total RNA was isolated from the tissues of male Sprague-Dawley rats (150-200 grams body weight; Charles River) or from the lactating mammary gland of female rats using a guanidium isothiocyanate-lithium chloride method. RNAs were fractionated on 1% agarose/formaldehyde gels with capillary transfer to nitrocellulose (MSI). cDNA probes were labelled by random priming.

Hybridization was carried in 5x Denhardt's, 0.2 M Tris (pH 7.4), 1M NaCl and 0.1 mg/ml salmon sperm DNA at 42°C for 20 hours. Filters were washed sequentially with 2X SSPE/0.1% SDS (room temperature; 2 x 15 minutes), 0.2 X SSPE/0.1% SDS (room temperature; 2 x 15 minutes) and with 0.2X SSPE/0.1% SDS (55°C; 2 x 15 minutes). Autoradiography on Kodak XAR film with enhancing screens was at -80°C for 18-48 hours.

#### *DNA Sequence Analysis and DNA sequences*

DNA sequences were analyzed using MacVector(r) and the GCG software package. Sequences were compared to the data base using BLAST and GCG; amino acid alignments were made using the Pileup program of GCG. Sequences were formatted using an Excel(r) macro. The DNA sequences described herein have been deposited in the GenBank with the following accession numbers: rat liver AMPK  $\beta$  (U42411), rat liver AMPK  $\gamma$  (U42413) and human fetal liver AMPK  $\gamma$  (U42412).

## SEQUENCE LISTING

## (1) GENERAL INFORMATION:

(i) APPLICANT: Dartmouth College, St. Vincent's  
Institute of Medical Research, Kemp et al.

(ii) TITLE OF INVENTION: Novel AMP Activated Protein  
Kinase

(iii) NUMBER OF SEQUENCES:

(iv) CORRESPONDENCE ADDRESS:

(A) ADDRESSEE: Jane Massey Licata, Esq.

(B) STREET: 210 Lake Drive East, Suite 201

(C) CITY: Cherry Hill

(D) STATE: NJ

(E) COUNTRY: USA

(F) ZIP: 08002

(v) COMPUTER READABLE FORM:

(A) MEDIUM TYPE: DISKETTE, 3.5 INCH, 1.44 Mb STORAGE

(B) COMPUTER: IBM 486

(C) OPERATING SYSTEM: WINDOWS FOR WORKGROUPS

(D) SOFTWARE: WORDPERFECT 5.1

(vi) CURRENT APPLICATION DATA:

(A) APPLICATION NUMBER: Not yet assigned

(B) FILING DATE: Herewith

(C) CLASSIFICATION:

(vii) PRIOR APPLICATION DATA:

(A) APPLICATION NUMBER: PN7450

(B) FILING DATE: 8 JAN 1996

(viii) ATTORNEY/AGENT INFORMATION:

(A) NAME: Jane Massey Licata

(B) REGISTRATION NUMBER: 32,257

- 30 -

(C) REFERENCE/DOCKET NUMBER: DC-0028

(ix) TELECOMMUNICATION INFORMATION:

(A) TELEPHONE: (609) 779-2400

(B) TELEFAX: (609) 779-8488

(2) INFORMATION FOR SEQ ID NO: 1:

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 345

(B) TYPE: Amino acid

(D) TOPOLOGY: Linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 1:

|     |     |     |     |     |     |     |     |     |     |     |     |     |     |     |
|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|
| MET | ALA | GLU | LYS | GLN | LYS | HIS | GLY | ARG | VAL | LYS | ILE | GLY | HIS | TYR |
| 1   |     |     |     | 5   |     |     |     |     | 10  |     |     |     |     | 15  |
| ILE | LEU | GLY | ASP | THR | LEU | GLY | VAL | GLY | THR | PHE | GLY | LYS | VAL | LYS |
|     |     |     |     | 20  |     |     |     |     | 25  |     |     |     |     | 30  |
| VAL | GLY | LYS | HIS | GLU | LEU | THR | GLY | HIS | LYS | VAL | ALA | VAL | LYS | ILE |
|     |     |     |     | 35  |     |     |     |     | 40  |     |     |     |     | 45  |
| LEU | ASN | ARG | GLN | LYS | ILE | ARG | LEU | ASP | VAL | VAL | GLY | LYS | ILE | ARG |
|     |     |     |     | 50  |     |     |     |     | 55  |     |     |     |     | 60  |
| ARG | GLU | ILE | GLN | ASN | LEU | LYS | LEU | PHE | ARG | HIS | PRO | HIS | ILE | ILE |
|     |     |     |     | 65  |     |     |     |     | 70  |     |     |     |     | 75  |
| LYS | LEU | TYR | GLN | VAL | ILE | SER | THR | PRO | SER | ASP | ILE | PHE | MET | VAL |
|     |     |     |     | 80  |     |     |     |     | 85  |     |     |     |     | 90  |
| MET | GLU | TYR | VAL | SER | GLY | GLY | GLU | LEU | PHE | ASP | TYR | ILE | CYS | LYS |
|     |     |     |     | 95  |     |     |     |     | 100 |     |     |     |     | 105 |
| ASN | GLY | ARG | LEU | ASP | GLU | LYS | GLU | SER | ARG | ARG | LEU | PHE | GLN | GLN |
|     |     |     |     | 110 |     |     |     |     | 115 |     |     |     |     | 120 |
| ILE | LEU | SER | GLY | VAL | ASP | TYR | CYS | HIS | ARG | HIS | MET | VAL | VAL | HIS |
|     |     |     |     | 125 |     |     |     |     | 130 |     |     |     |     | 135 |
| ARG | ASP | LEU | LYS | PRO | GLU | ASN | VAL | LEU | LEU | ASP | ALA | HIS | MET | ASN |
|     |     |     |     | 140 |     |     |     |     | 145 |     |     |     |     | 150 |
| ALA | LYS | ILE | ALA | ASP | PHE | GLY | LEU | SER | ASN | MET | MET | SER | ASP | GLY |
|     |     |     |     | 155 |     |     |     |     | 160 |     |     |     |     | 165 |
| GLU | PHE | LEU | ARG | THR | SER | CYS | GLY | SER | PRO | ASN | TYR | ALA | ALA | PRO |
|     |     |     |     | 170 |     |     |     |     | 175 |     |     |     |     | 180 |
| GLU | VAL | ILE | SER | GLY | ARG | LEU | TYR | ALA | GLY | PRO | GLU | VAL | ASP | ILE |
|     |     |     |     | 185 |     |     |     |     | 190 |     |     |     |     | 195 |
| TRP | SER | SER | GLY | VAL | ILE | LEU | TYR | ALA | LEU | LEU | CYS | GLY | THR | LEU |
|     |     |     |     | 200 |     |     |     |     | 205 |     |     |     |     | 210 |
| PRO | PHE | ASP | ASP | ASP | HIS | VAL | PRO | THR | LEU | PHE | LYS | LYS | ILE | CYS |
|     |     |     |     | 215 |     |     |     |     | 220 |     |     |     |     | 225 |
| ASP | GLY | ILE | PHE | TYR | THR | PRO | GLN | TYR | LEU | ASN | PRO | SER | VAL | ILE |
|     |     |     |     | 230 |     |     |     |     | 235 |     |     |     |     | 240 |
| SER | LEU | LEU | LYS | HIS | MET | LEU | GLN | VAL | ASP | PRO | MET | LYS | ARG | ALA |
|     |     |     |     | 245 |     |     |     |     | 250 |     |     |     |     | 255 |
| THR | ILE | LYS | ASP | ILE | ARG | GLU | HIS | GLU | TRP | PHE | LYS | GLN | ASP | LEU |
|     |     |     |     | 260 |     |     |     |     | 265 |     |     |     |     | 270 |



- 31 -

|     |     |     |     |     |     |     |     |     |     |     |     |     |     |     |
|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|
| PRO | LYS | TYR | LEU | PHE | PRO | GLU | ASP | PRO | SER | TYR | SER | SER | THR | MET |
|     |     |     |     | 275 |     |     |     |     | 280 |     |     |     |     | 285 |
| ILE | ASP | ASP | GLU | ALA | LEU | LYS | GLU | VAL | CYS | GLU | LYS | PHE | GLU | CYS |
|     |     |     |     | 290 |     |     |     |     | 295 |     |     |     |     | 300 |
| SER | GLU | GLU | GLU | VAL | LEU | SER | CYS | LEU | TYR | ASN | ARG | ASN | HIS | GLN |
|     |     |     |     | 305 |     |     |     |     | 310 |     |     |     |     | 315 |
| ASP | PRO | LEU | ALA | VAL | ALA | TYR | HIS | LEU | ILE | ILE | ASP | ASN | ARG | ARG |
|     |     |     |     | 320 |     |     |     |     | 325 |     |     |     |     | 330 |
| ILE | MET | ASN | GLU | ALA | LYS | ASP | PHE | TYR | LEU | ALA | THR | SER | PRO | PRO |
|     |     |     |     | 335 |     |     |     |     | 340 |     |     |     |     | 345 |

## (2) INFORMATION FOR SEQ ID NO: 2:

## (i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 11

(B) TYPE: Amino acid

(D) TOPOLOGY: Linear

## (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 2:

ASP SER PHE LEU ASP ASP HIS HIS LEU THR ARG

1

5

10

## (2) INFORMATION FOR SEQ ID NO: 3:

## (i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 56

(B) TYPE: Amino acid

(D) TOPOLOGY: Linear

## (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 3:

|     |     |     |     |     |     |     |     |     |     |     |     |     |     |     |
|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|
| PRO | HIS | PRO | GLU | ARG | VAL | PRO | PHE | LEU | VAL | ALA | GLU | THR | PRO | ARG |
| 1   |     |     |     | 5   |     |     |     |     | 10  |     |     |     |     | 15  |
| ALA | ARG | HIS | THR | LEU | ASP | GLU | LEU | ASN | PRO | GLN | LYS | SER | LYS | HIS |
|     |     |     |     | 20  |     |     |     |     | 25  |     |     |     |     | 30  |
| GLN | GLY | VAL | ARG | LYS | ALA | LYS | TRP | HIS | LEU | GLY | ILE | ARG | SER | GLN |
|     |     |     |     | 35  |     |     |     |     | 40  |     |     |     |     | 45  |
| SER | ARG | PRO | ASN | ASP | ILE | MET | ALA | GLU | VAL | CYS |     |     |     |     |
|     |     |     |     | 50  |     |     |     |     | 55  |     |     |     |     |     |

## (2) INFORMATION FOR SEQ ID NO: 4:

## (i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 70

(B) TYPE: Amino acid

- 32 -

(D) TOPOLOGY: Linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 4:

|     |     |     |     |     |     |     |     |     |     |     |     |     |     |     |
|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|
| ARG | ALA | ILE | LYS | GLN | LEU | ASP | TYR | GLU | TRP | LYS | VAL | VAL | ASN | PRO |
| 1   |     |     |     | 5   |     |     |     |     | 10  |     |     |     |     | 15  |
| TYR | TYR | LEU | ARG | VAL | ARG | ARG | LYS | ASN | PRO | VAL | THR | SER | THR | PHE |
|     |     |     |     | 20  |     |     |     |     | 25  |     |     |     |     | 30  |
| SER | LYS | MET | SER | LEU | GLN | LEU | TYR | GLN | VAL | ASP | SER | ARG | THR | TYR |
|     |     |     |     | 35  |     |     |     |     | 40  |     |     |     |     | 45  |
| LEU | LEU | ASP | PHE | ARG | SER | ILE | ASP | ASP | GLU | ILE | THR | GLU | ALA | LYS |
|     |     |     |     | 50  |     |     |     |     | 55  |     |     |     |     | 60  |
| SER | GLY | THR | ALA | THR | PRO | GLN | ARG | SER | GLY |     |     |     |     |     |
|     |     |     |     | 67  |     |     |     |     | 70  |     |     |     |     |     |

(2) INFORMATION FOR SEQ ID NO: 5:

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 64

(B) TYPE: Amino acid

(D) TOPOLOGY: Linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 5:

|     |     |     |     |     |     |     |     |     |     |     |     |     |     |     |
|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|
| SER | ILE | SER | ASN | TYR | ARG | SER | CYS | GLN | ARG | SER | ASP | SER | ASP | ALA |
| 1   |     |     |     | 5   |     |     |     |     | 10  |     |     |     |     | 15  |
| GLU | ALA | GLN | GLY | LYS | PRO | SER | GLU | VAL | SER | LEU | THR | SER | SER | VAL |
|     |     |     |     | 20  |     |     |     |     | 25  |     |     |     |     | 30  |
| THR | SER | LEU | ASP | SER | SER | PRO | VAL | ASP | VAL | ALA | PRO | ARG | PRO | GLY |
|     |     |     |     | 35  |     |     |     |     | 40  |     |     |     |     | 45  |
| SER | HIS | THR | ILE | GLU | PHE | PHE | GLU | MET | CYS | ALA | ASN | LEU | ILE | LYS |
|     |     |     |     | 50  |     |     |     |     | 55  |     |     |     |     | 60  |
| ILE | LEU | ALA | GLN |     |     |     |     |     |     |     |     |     |     |     |

(2) INFORMATION FOR SEQ ID NO: 6:

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 242

(B) TYPE: Amino acid

(D) TOPOLOGY: Linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 6:

|     |     |     |     |     |     |     |     |     |     |     |     |     |     |     |
|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|
| GLY | HIS | TYR | ILE | LEU | GLY | ASP | THR | LEU | GLY | VAL | GLY | THR | PHE | GLY |
| 1   |     |     |     | 5   |     |     |     |     | 10  |     |     |     |     | 15  |
| LYS | VAL | LYS | VAL | GLY | LYS | HIS | GLU | LEU | THR | GLY | HIS | LYS | VAL | ALA |
|     |     |     |     | 20  |     |     |     |     | 25  |     |     |     |     | 30  |
| VAL | LYS | ILE | LEU | ASN | ARG | GLN | LYS | ILE | ARG | SER | LEU | ASP | VAL | VAL |
|     |     |     |     | 35  |     |     |     |     | 40  |     |     |     |     | 45  |
| GLY | LYS | ILE | ARG | ARG | GLU | ILE | GLN | ASN | LEU | LYS | LEU | PHE | ARG | HIS |

- 33 -

|     |     |     |     |     |     |     |     |     |     |     |     |     |     |     |     |
|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|
| PRO | HIS | ILE | ILE | LYS | LEU | TYR | GLN | VAL | ILE | SER | THR | PRO | SER | ASP | 60  |
|     |     |     |     | 65  |     |     |     |     | 70  |     |     |     |     |     | 75  |
| ILE | PHE | MET | VAL | MET | GLU | TYR | VAL | SER | GLY | GLY | GLU | LEU | PHE | ASP | 90  |
|     |     |     |     | 80  |     |     |     |     | 85  |     |     |     |     |     | 90  |
| TYR | ILE | CYS | LYS | ASN | GLY | ARG | LEU | ASP | GLU | LYS | GLU | SER | ARG | ARG | 105 |
|     |     |     |     | 95  |     |     |     |     | 100 |     |     |     |     |     | 105 |
| LEU | PHE | GLN | GLN | ILE | LEU | SER | GLY | VAL | ASP | TYR | CYS | HIS | ARG | HIS | 120 |
|     |     |     |     | 110 |     |     |     |     | 115 |     |     |     |     |     | 120 |
| MET | VAL | VAL | HIS | ARG | ASP | LEU | LYS | PRO | GLU | ASN | VAL | LEU | LEU | ASP |     |
|     |     |     |     |     |     |     |     |     |     |     |     |     |     |     |     |
| ALA | HIS | MET | ASN | ALA | LYS | ILE | ALA | ASP | PHE | GLY | LEU | SER | ASN | MET |     |
|     |     |     |     | 125 |     |     |     |     | 130 |     |     |     |     |     | 135 |
| MET | SER | ASP | GLY | GLU | PHE | LEU | ARG | THR | SER | CYS | GLY | SER | PRO | ASN |     |
|     |     |     |     | 140 |     |     |     |     | 145 |     |     |     |     |     | 150 |
| TYR | ALA | ALA | PRO | GLU | VAL | ILE | SER | GLY | ARG | LEU | TYR | ALA | GLY | PRO |     |
|     |     |     |     | 155 |     |     |     |     | 160 |     |     |     |     |     | 165 |
| GLU | VAL | ASP | ILE | TRP | SER | SER | GLY | VAL | ILE | LEU | TYR | ALA | LEU | LEU |     |
|     |     |     |     | 170 |     |     |     |     | 175 |     |     |     |     |     | 180 |
| CYS | GLY | THR | LEU | PRO | PHE | ASP | ASP | ASP | HIS | VAL | PRO | THR | LEU | PHE |     |
|     |     |     |     | 185 |     |     |     |     | 190 |     |     |     |     |     | 195 |
| LYS | LYS | ILE | CYS | ASP | GLY | ILE | PHE | TYR | THR | PRO | GLN | TYR | LEU | ASN |     |
|     |     |     |     | 200 |     |     |     |     | 205 |     |     |     |     |     | 210 |
| PRO | SER | VAL | ILE | SER | LEU | LEU | LYS | HIS | MET | LEU | GLN | VAL | ASP | PRO |     |
|     |     |     |     | 215 |     |     |     |     | 220 |     |     |     |     |     | 225 |
| MET | LYS | ARG | ALA | THR | ILE | LYS | ASP | ILE | ARG | GLU | HIS | GLU | TRP | PHE |     |
|     |     |     |     | 230 |     |     |     |     | 235 |     |     |     |     |     | 240 |
| LYS | GLN |     |     |     |     |     |     |     |     |     |     |     |     |     |     |

(2) INFORMATION FOR SEQ ID NO: 7:

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 55

(B) TYPE: Amino acid

(D) TOPOLOGY: Linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 7:

|     |     |     |     |     |     |     |     |     |     |     |     |     |     |     |  |
|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|--|
| GLU | ALA | LEU | LYS | GLU | VAL | CYS | GLU | LYS | PHE | GLU | CYS | SER | GLU | GLU |  |
| 1   |     |     |     | 5   |     |     |     |     | 10  |     |     |     |     | 15  |  |
| GLU | VAL | LEU | SER | CYS | LEU | TYR | ASN | ARG | ASN | HIS | GLN | ASP | PRO | LEU |  |
|     |     |     |     | 20  |     |     |     |     | 25  |     |     |     |     | 30  |  |
| ALA | VAL | ALA | TYR | HIS | LEU | ILE | ILE | ASP | ASN | ARG | ARG | ILE | MET | ASN |  |
|     |     |     |     | 35  |     |     |     |     | 40  |     |     |     |     | 45  |  |
| GLU | ALA | LYS | ASP | PHE | TYR | LEU | ALA | THR | SER |     |     |     |     |     |  |
|     |     |     |     | 50  |     |     |     |     | 55  |     |     |     |     |     |  |

(2) INFORMATION FOR SEQ ID NO: 8:

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 9

- 34 -

(B) TYPE: Amino acid

(D) TOPOLOGY: Linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 8:

PHE LEU ASP ASP HIS HIS LEU THR ARG

1

5

(2) INFORMATION FOR SEQ ID NO: 9:

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 56

(B) TYPE: Amino acid

(D) TOPOLOGY: Linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 9:

|     |     |     |     |     |     |     |     |     |     |     |     |     |     |     |
|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|
| PRO | HIS | PRO | GLU | ARG | VAL | PRO | PHE | LEU | VAL | ALA | GLU | THR | PRO | ARG |
| 1   |     |     |     | 5   |     |     |     |     | 10  |     |     |     |     | 15  |
| ALA | ARG | HIS | THR | LEU | ASP | GLU | LEU | ASN | PRO | GLN | LYS | SER | LYS | HIS |
|     |     |     |     | 20  |     |     |     |     | 25  |     |     |     |     | 30  |
| GLN | GLY | VAL | ARG | LYS | ALA | LYS | TRP | HIS | LEU | GLY | ILE | ARG | SER | GLN |
|     |     |     |     | 35  |     |     |     |     | 40  |     |     |     |     | 45  |
| SER | ARG | PRO | ASN | ASP | ILE | MET | ALA | GLU | VAL | CYS |     |     |     |     |
|     |     |     |     | 50  |     |     |     |     | 55  |     |     |     |     |     |

(2) INFORMATION FOR SEQ ID NO: 10:

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 10

(B) TYPE: Amino acid

(D) TOPOLOGY: Linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 10:

ARG PRO ASN ASP ILE MET ALA GLU VAL CYS

1

5

10

(2) INFORMATION FOR SEQ ID NO: 11:

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 70

(B) TYPE: Amino acid

(D) TOPOLOGY: Linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 11:

|     |     |     |     |     |     |     |     |     |     |     |     |     |     |     |
|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|
| ARG | ALA | ILE | LYS | GLN | LEU | ASP | TYR | GLU | TRP | LYS | VAL | VAL | ASN | PRO |
| 1   |     |     |     | 5   |     |     |     |     | 10  |     |     |     |     | 15  |
| TYR | TYR | LEU | ARG | VAL | ARG | ARG | LYS | ASN | PRO | VAL | THR | SER | THR | TYR |
|     |     |     |     | 20  |     |     |     |     | 25  |     |     |     |     | 30  |
| SER | LYS | MET | SER | LEU | GLN | LEU | TYR | GLN | VAL | ASP | SER | ARG | THR | TYR |
|     |     |     |     | 35  |     |     |     |     | 40  |     |     |     |     | 45  |
| LEU | LEU | ASP | PHE | ARG | SER | ILE | ASP | ASP | GLU | ILE | THR | GLU | ALA | LYS |
|     |     |     |     | 50  |     |     |     |     | 55  |     |     |     |     | 60  |
| SER | GLY | THR | ALA | THR | PRO | GLN | ARG | SER | GLY |     |     |     |     |     |
|     |     |     |     | 65  |     |     |     |     | 70  |     |     |     |     |     |

(2) INFORMATION FOR SEQ ID NO: 12:

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 64

(B) TYPE: Amino acid

(D) TOPOLOGY: Linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 12:

[illegible]

(2) INFORMATION FOR SEQ ID NO: 13:

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 22

(B) TYPE: Amino acid

(D) TOPOLOGY: Linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 13:

```

ASP GLY ARG VAL LYS ILE GLY HIS TYR ILE LEU GLY ASP THR LEU
1           5           10           15
GLY VAL GLY THR PHE GLY LYS
                20

```

(2) INFORMATION FOR SEQ ID NO: 14:

- 36 -

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 16

(B) TYPE: Amino acid

(D) TOPOLOGY: Linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 14:

ASP GLU LYS GLU SER ARG ARG LEU PHE GLN GLN ILE LEU SER GLY

1                      5                      10                      15

VAL

(2) INFORMATION FOR SEQ ID NO: 15:

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 33

(B) TYPE: Amino acid

(D) TOPOLOGY: Linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 15:

ASP LEU LYS PRO GLU ASN VAL LEU LEU ASP ALA HIS MET ASN ALA

```

1          5          10          15
LYS ILE ALA ASP PHE GLY LEU SER ASN MET MET SER ASP GLY GLU
          20          25          30

```

PHE LEU ARG

(2) INFORMATION FOR SEQ ID NO: 16:

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 13

(B) TYPE: Amino acid

(D) TOPOLOGY: Linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 16:

GLU VAL ILE SER GLY ARG LEU TYR ALA GLY PRO GLU VAL

1                      5                      10

(2) INFORMATION FOR SEQ ID NO: 17:

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 9

- 37 -

(B) TYPE: Amino acid

(D) TOPOLOGY: Linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 17:

XAA MET LEU GLN VAL ASP PRO MET LYS

1

5

(2) INFORMATION FOR SEQ ID NO: 18:

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 35

(B) TYPE: Amino acid

(D) TOPOLOGY: Linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 18:

|     |     |     |     |     |     |     |     |     |     |     |     |     |     |     |
|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|
| LYS | ASP | ILE | ARG | GLU | HIS | GLU | XAA | PHE | LYS | GLN | ASP | LEU | PRO | LYS |
| 1   |     |     |     | 5   |     |     |     |     | 10  |     |     |     |     | 15  |
| TYR | LEU | PHE | PRO | GLU | ASP | PRO | SER | TYR | SER | XAA | THR | MET | ILE | ASP |
|     |     |     |     | 20  |     |     |     |     | 25  |     |     |     |     | 30  |
| ASP | GLU | ALA | LEU | LYS |     |     |     |     |     |     |     |     |     |     |
|     |     |     |     | 35  |     |     |     |     |     |     |     |     |     |     |

(2) INFORMATION FOR SEQ ID NO: 19:

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 17

(B) TYPE: Amino acid

(D) TOPOLOGY: Linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 19:

|     |     |     |     |     |     |     |     |     |     |     |     |     |     |     |
|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|
| XAA | XAA | GLN | ASP | PRO | LEU | ALA | VAL | ALA | TYR | HIS | LEU | ILE | ILE | ASP |
| 1   |     |     |     | 5   |     |     |     |     | 10  |     |     |     |     | 15  |
| ASN | ARG |     |     |     |     |     |     |     |     |     |     |     |     |     |

(2) INFORMATION FOR SEQ ID NO: 20:

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 9

(B) TYPE: Amino acid

(D) TOPOLOGY: Linear

- 38 -

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 20:

ASP PHE TYR LEU ALA THR SER PRO PRO

1 5

(2) INFORMATION FOR SEQ ID NO: 21:

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 11

(B) TYPE: Amino acid

(D) TOPOLOGY: Linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 21:

ASP SER PHE LEU ASP ASP HIS HIS LEU THR ARG

1 5 10

(2) INFORMATION FOR SEQ ID NO: 22:

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 10

(B) TYPE: Amino acid

(D) TOPOLOGY: Linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 22:

VAL PRO PHE LEU VAL ALA GLU THR PRO ARG

1 5 10

(2) INFORMATION FOR SEQ ID NO: 23:

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 23

(B) TYPE: Amino acid

(D) TOPOLOGY: Linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 23:

ASP GLU LEU ASN PRO GLN LYS XAA LYS HIS GLN GLY VAL ARG LYS

1 5 10 15

ALA LYS XAA HIS LEU GLY ILE ARG



(2) INFORMATION FOR SEQ ID NO: 24:

(A) LENGTH: 19

(B) TYPE: Amino acid

(D) TOPOLOGY: Linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 24:

GLN LEU ASP TYR GLU XAA LYS VAL VAL ASN PRO TYR TYR LEU ARG

|   |   |    |    |
|---|---|----|----|
| 1 | 5 | 10 | 15 |
|---|---|----|----|

VAL ARG ARG LYS

## (2) INFORMATION FOR SEQ ID NO: 25:

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 25

(B) TYPE: Amino acid

(D) TOPOLOGY: Linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 25:

LYS MET SER LEU GLN LEU TYR GLN VAL ASP SER ARG THR TYR LEU

|   |   |    |    |
|---|---|----|----|
| 1 | 5 | 10 | 15 |
|---|---|----|----|

LEU ASP PHE ARG SER ILE ASP ASP XAA ILE

20 25

(2) INFORMATION FOR SEQ ID NO: 26:

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 18

(B) TYPE: Amino acid

(D) TOPOLOGY: Linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 26:

ASP ALA GLU ALA GLN GLY LYS SER SER GLU ALA SER LEU THR XAA

1                      5                      10                      15

SER VAL THR

(2) INFORMATION FOR SEQ ID NO: 27:

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 17

(B) TYPE: Amino acid

(D) TOPOLOGY: Linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 27:

ILE GLY HIS TYR ILE LEU GLY ASP THR LEU GLY VAL GLY THR PHE

1 5 10 15

GLY LYS

(2) INFORMATION FOR SEQ ID NO: 28:

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 26

(B) TYPE: Amino acid

(D) TOPOLOGY: Linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 28:

LEU TYR GLN VAL ILE SER THR PRO SER ASP ILE PHE MET VAL MET

1 5 10 15

GLU TYR VAL SER GLY GLY GLU LEU PHE ASP TYR

20 25

(2) INFORMATION FOR SEQ ID NO: 29:

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 12

(B) TYPE: Amino acid

(D) TOPOLOGY: Linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 29:

ARG LEU PHE GLN GLN ILE LEU SER GLY VAL ASP TYR

1 5 10

- 41 -

## (2) INFORMATION FOR SEQ ID NO: 30:

## (i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 11

(B) TYPE: Amino acid

(D) TOPOLOGY: Linear

## (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 30:

ASP LEU LYS PRO GLU ASN VAL LEU LEU ASP ALA

1 5 10

## (2) INFORMATION FOR SEQ ID NO: 31:

## (i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 17

(B) TYPE: Amino acid

(D) TOPOLOGY: Linear

## (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 31:

ILE ALA ASP PHE GLY LEU SER ASN MET MET SER ASP GLY GLU PHE

1 5 10 15

LEU ARG

## (2) INFORMATION FOR SEQ ID NO: 32:

## (i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 22

(B) TYPE: Amino acid

(D) TOPOLOGY: Linear

## (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 32:

LYS ILE XAA ASP GLY ILE PHE TYR THR PRO GLN TYR LEU ASN PRO

1 5 10 15

XAA VAL ILE XAA LEU LEU LYS

20

## (2) INFORMATION FOR SEQ ID NO: 33:

- 42 -

## (i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 5  
(B) TYPE: Amino acid  
(D) TOPOLOGY: Linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 33:

ASP ILE ARG GLU HIS

1 5

## (2) INFORMATION FOR SEQ ID NO: 34:

## (i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 20  
(B) TYPE: Amino acid  
(D) TOPOLOGY: Linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 34:

TYR LEU PHE PRO GLU ASP PRO SER TYR SER XAA XAA MET ILE ASP

1 5 10 15

ASP GLU ALA LEU LYS

20

## (2) INFORMATION FOR SEQ ID NO: 35:

## (i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 16  
(B) TYPE: Amino acid  
(D) TOPOLOGY: Linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 35:

ASN HIS GLN ASP PRO LEU ALA VAL ALA TYR HIS LEU ILE ILE ASP

1 5 10 15

ASN

## (2) INFORMATION FOR SEQ ID NO: 36:

## (i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 9  
(B) TYPE: Amino acid  
(D) TOPOLOGY: Linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 36:

ASP PHE TYR LEU ALA THR XAA PRO PRO

1

5

(2) INFORMATION FOR SEQ ID NO: 37:

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 9  
(B) TYPE: Amino acid  
(D) TOPOLOGY: Linear

(xi) SEQUENCE DESCRIPTION: SEQ ID. NO: 37:

ASP XAA PHE LEU ASP ASP HIS XAA LEU

1

5

(2) INFORMATION FOR SEQ ID NO: 38:

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 10  
(B) TYPE: Amino acid  
(D) TOPOLOGY: Linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 38:

VAL PRO PHE LEU VAL ALA GLU THR PRO ARG

1

5

10

(2) INFORMATION FOR SEQ ID NO: 39:

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 5  
(B) TYPE: Amino acid  
(D) TOPOLOGY: Linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 39:

- 44 -

TRP HIS LEU GLY ILE

1 5

(2) INFORMATION FOR SEQ ID NO: 40:

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 14

(B) TYPE: Amino acid

(D) TOPOLOGY: Linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 40:

XAA GLN SER ARG PRO ASN ASP ILE MET ALA GLU VAL XAA ARG

1 5 10

(2) INFORMATION FOR SEQ ID NO: 41:

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 10

(B) TYPE: Amino acid

(D) TOPOLOGY: Linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 41:

VAL VAL ASN PRO TYR TYR LEU ARG VAL ARG

1 5 10

(2) INFORMATION FOR SEQ ID NO: 42:

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 18

(B) TYPE: Amino acid

(D) TOPOLOGY: Linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 42:

MET SER LEU GLN LEU TYR GLN VAL ASP SER ARG THR TYR LEU LEU

1 5 10 15

LEU PHE ARG

(2) INFORMATION FOR SEQ ID NO: 43:

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 12

(B) TYPE: Amino acid

(D) TOPOLOGY: Linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 43:

XAA ASP SER ASP ALA GLU ALA GLN GLY LYS PRO SER

1

5

10

(2) INFORMATION FOR SEQ ID NO: 44:

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 1647

(B) TYPE: Nucleic acid

(C) STRANDEDNESS: Single

(D) TOPOLOGY: Linear

(iv) ANTI-SENSE: No

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 44:

|             |            |            |             |            |      |
|-------------|------------|------------|-------------|------------|------|
| ATTGCCCGAGA | AGCAGAAGCA | CGACGGGCGG | GTGAAGATCG  | GCCACTACAT | 50   |
| CCTGGGGGAC  | ACGCTGGGCG | TCGGCACCTT | CGGGAAAGTG  | AAGGTGGGCA | 100  |
| AGCACGAGTT  | GACTGGACAT | AAAGTTGCTG | TGAAGATACT  | CAACCGGCAG | 150  |
| AAGATTCGAA  | GCCTGGACGT | GGTCGGGAAA | ATCCGCAGAG  | AGATCCAGAA | 200  |
| CCTGAAGCTT  | TTCAGGCACC | CTCATATAAT | CAAAC TGTAC | CAGGTCATCA | 250  |
| GTACACCGTC  | TGATATTTTC | ATGGTCATGG | AATATGTCTC  | AGGAGGAGAG | 300  |
| CTATTTGATT  | ATATCTGTAA | AAATGGAAGG | TTGGACGAAA  | AGGAGAGTCG | 350  |
| ACGTCTGTTC  | CAGCAGATCC | TTTCTGGTGT | GGACTATTGT  | CACAGGCATA | 400  |
| TGGTGGTCCA  | CAGAGATTTG | AAACCTGAAA | ACGTCCTGCT  | TGATGCACAC | 450  |
| ATGAATGCAA  | AGATAGCCGA | CTTCGGTCTT | TCAAACATGA  | TGTCAGATGG | 500  |
| TGAATTTTTA  | AGAACGAGCT | GTGGCTCGCC | CAATTATGCT  | GCACCAGAAG | 550  |
| TAATTTTCAGG | AAGATTCTAC | GCAGGCCCTG | AAGTAGACAT  | CTGGAGCAGC | 600  |
| GGGGTCATTC  | TCTATGCTTT | GCTGTGTGGA | ACTCTCCCTT  | TTGATGATGA | 650  |
| CCACGTGCCA  | ACTCTTTTTA | AGAAGATATG | TGACGGGATA  | TTTTATACCC | 700  |
| CTCAGTATTT  | GAATCCCTCT | GTAATAAGCC | TTTTGAAGCA  | TATGCTGCAG | 750  |
| GTAGATCCTA  | TGAAGAGGGC | CACAATAAAA | GATATCAGGG  | AACATGAATG | 800  |
| GTTTAAGCAG  | GACCTTCCAA | AATATCTCTT | TCCTGAAGAC  | CCGTCTTATA | 850  |
| GTTCAACCAT  | GATTGATGAT | GAAGCCTTAA | AAGAAGTGTG  | TGAGAAGTTC | 900  |
| GAGTGCTCAG  | AGGAGGAGGT | CCTCAGCTGC | CTGTACAACA  | GAAACCACCA | 950  |
| GGACCCACTG  | GCAGTTGCCT | ACCACCTCAT | AATAGACAAC  | AGGAGAATAA | 1000 |
| TGAACGAAGC  | CAAAGATTTT | TACTTGGCAA | CAAGCCCACC  | CGATTCTTTT | 1050 |
| CTCGATGATC  | ACCATTTAAC | TCCGCCTCAC | CCTGAGAGAG  | TACCATTCTT | 1100 |
| GTTTGCCGAA  | ACACCAAGGG | C GGACACAC | CCTAGATGAA  | TTAAACCCAT | 1150 |
| AGAAATCCAA  | ACACCAAGGC | GTACGGAAGG | CAAAGTGGCA  | TTTGGGGATT | 1200 |
| CGAAGTCAAA  | GCCGACCCAA | TGACATCATG | GCAGAAGTGT  | GTAGAGCAAT | 1250 |
| CAAGCAGTTG  | GACTATGAAT | GGAAGGTTGT | AAACCCCTAT  | TATTTGCGTG | 1300 |

- 46 -

TGCGAAGGAA GAACCCTGTG ACAAGCACAT TTTCCAAAAT GAGTCTACAG 1350  
CTATACCAAG TGGATAGTAG GACTTACTTA TTGGATTTCC GAAGTATTGA 1400  
TGATGAGATT ACAGAAGCCA AATCAGGGAC TGCTACTCCA CAGAGATCGG 1450  
GATCCATCAG CAACTATCGA TCTTGCCAAA GGAGCGACTC CGACGCCGAG 1500  
GCTCAAGGAA AGCCCTCAGA AGTCTCTCTT ACCTCATCCG TGACCTCCCT 1550  
CGACTCCTCT CCTGTTGACG TAGCTCCAAG ACCAGGAAGT CACACGATAG 1600  
AATTTTTTGA AATGTGTGCA AATCTAATTA AAATTCTTGC ACAGTAA 1647

## (2) INFORMATION FOR SEQ ID NO: 45:

## (i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 20

(B) TYPE: Amino Acid

(D) TOPOLOGY: Linear

## (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 45:

ASP PHE TYR LEU ALA THR SER PRO PRO ASP SER PHE LEU ASP ASP

1 5 10 15

HIS HIS LEU THR ARG

20

## (2) INFORMATION FOR SEQ ID NO: 46:

## (i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 12

(B) TYPE: Amino Acid

(D) TOPOLOGY: Linear

## (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 46:

PRO LEU SER ARG THR LEU SER VAL ALA ALA LYS LYS

1 5 10

## (2) INFORMATION FOR SEQ ID NO: 47:

## (i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 13

(B) TYPE: Amino Acid

(D) TOPOLOGY: Linear

## (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 47:



- 47 -

LEU LYS LYS LEU THR LEU ARG ALA SER PHE SER ALA GLN

1 5 10

(2) INFORMATION FOR SEQ ID NO: 48:

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 14

(B) TYPE: Amino Acid

(D) TOPOLOGY: Linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 48:

PRO GLY LEU LYS PRO HIS PRO GLU ARG MET PRO PRO LEU ILE

1 5 10

(2) INFORMATION FOR SEQ ID NO: 49:

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 15

(B) TYPE: Amino Acid

(D) TOPOLOGY: Linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 49:

HIS MET ARG SER ALA MET SER GLY LEU HIS LEU VAL LYS ARG ARG

1 5 10 15

(2) INFORMATION FOR SEQ ID NO: 50:

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 20

(B) TYPE: Amino Acid

(D) TOPOLOGY: Linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 50:

ASP PHE TYR LEU ALA THR SER PRO PRO ASP SER PHE LEU ASP ASP

1 5 10 15

HIS HIS LEU THR ARG

- 48 -

## (2) INFORMATION FOR SEQ ID NO: 51:

## (i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 15  
(B) TYPE: Amino Acid  
(D) TOPOLOGY: Linear

## (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 51:

MET ASP ASP SER ALA MET HIS ILE PRO PRO GLY LEU LYS PRO HIS  
1 5 10 15

## (2) INFORMATION FOR SEQ ID NO: 52:

## (i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 17  
(B) TYPE: Amino Acid  
(D) TOPOLOGY: Linear

## (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 52:

VAL VAL ASP ILE TYR SER LYS PHE ASP VAL ILE ASN LEU ALA ALA  
1 5 10 15  
GLU LYS

## (2) INFORMATION FOR SEQ ID NO: 53:

## (i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 19  
(B) TYPE: Amino Acid  
(D) TOPOLOGY: Linear

## (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 53:

GLY CYS GLY GLY ALA THR CYS CYS GLY THR ASN GLY ALA TYR ALA  
1 5 10 15  
THR HIS THR ALA

## (2) INFORMATION FOR SEQ ID NO: 54:

## (i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 22

(B) TYPE: Amino Acid

(D) TOPOLOGY: Linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 54:

CYS GLY GLY ALA ALA THR THR CYS TYR THR THR TYR THR CYS ASN

1                      5                      10                      15

GLY CYS ASN GLY CYS ASN ALA

20

(2) INFORMATION FOR SEQ ID NO: 55:

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 8

(B) TYPE: Amino Acid

(D) TOPOLOGY: Linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 55:

LYS PHE ASP VAL ILE ASN LEU ALA

1 5

(2) INFORMATION FOR SEQ ID NO: 56:

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 23

(B) TYPE: Amino Acid

(D) TOPOLOGY: Linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 56:

ALA ALA ARG THR THR TYR GLY ALA TYR GLY THR ASN ALA THR HIS

1                      5                      10                      15

ALA ALA TYR CYS THR ASN GLY CYS

20

(2) INFORMATION FOR SEQ ID NO: 57:

(i) SEQUENCE CHARACTERISTICS:

- 50 -

- (A) LENGTH: 23  
(B) TYPE: Amino Acid  
(D) TOPOLOGY: Linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 57:

ALA ALA ARG THR THR TYR GLY ALA TYR GLY THR ASN ALA THR HIS  
1 5 10 15  
ALA ALA TYR THR THR ARG GLY CYS  
20

(2) INFORMATION FOR SEQ ID NO: 58:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 23  
(B) TYPE: Amino Acid  
(D) TOPOLOGY: Linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 58:

CYS THR CYS CYS ALA ALA GLY THR THR THR GLY ALA THR GLY THR  
1 5 10 15  
THR ALA THR CYS ALA ALA CYS CYS  
20

(2) INFORMATION FOR SEQ ID NO: 59:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 6  
(B) TYPE: Amino Acid  
(D) TOPOLOGY: Linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 59:

GLU GLU LEU GLN ILE GLY  
1 5

(2) INFORMATION FOR SEQ ID NO: 60:

(i) SEQUENCE CHARACTERISTICS:

- 51 -

- (A) LENGTH: 7  
 (B) TYPE: Amino Acid  
 (D) TOPOLOGY: Linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 60:

PHE PRO LYS PRO GLU PHE MET

1

5

(2) INFORMATION FOR SEQ ID NO: 61:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 1978  
 (B) TYPE: Nucleic acid  
 (C) STRANDEDNESS: Single  
 (D) TOPOLOGY: Linear

(iv) ANTI-SENSE: No

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 61:

|            |            |            |            |            |      |
|------------|------------|------------|------------|------------|------|
| CTCGCTGCGG | TCCAAGCAGG | TAAAGCGGGG | CTCGGCGAAC | GCGCGCGACC | 50   |
| CGAGGGGCGT | GGTCCGCGGT | CCCGGGGGTC | CCGGCCCGGC | CCTTCCCGCT | 100  |
| TCCCTGTGTC | CCCGCAGACA | CTTCGCCATG | GGCAATACGA | GCAGCGAGCG | 150  |
| CGCCGCGCTG | GAGCGGCAGG | CTGGCCATAA | GACGCCGCGG | AGGGACAGCT | 200  |
| CGGAGGGCAC | CAAGGATGGG | GACAGGCCCA | AGATCCTGAT | GGACAGCCCC | 250  |
| GAAGACGCCG | ACATCTTCCA | CACCGAGGAA | ATGAAGGCTC | CAGAGAAGGA | 300  |
| GGAGTTCTTG | GCGTGGCAGC | ACGACCTCGA | GGTGAATGAG | AAAGCCCCCG | 350  |
| CCCAGGCTCG | GCCCACCGTA | TTTCGATGGA | CAGGGGGTGG | AAAGGAGGTC | 400  |
| TACTTGTCTG | GATCCTTCAA | CAACTGGAGC | AAATTGCCCC | TCACTAGAAG | 450  |
| CCAAAACAAC | TTCGTAGCCA | TCCTGGACCT | NCCGGAAGGA | GAGCATCAGT | 500  |
| ACAAGTTCTT | TGTGGATGGC | CAGTGGACCC | ACGATCCTTC | CGAGCCAATA | 550  |
| GTAACCAGCC | AGCTTGGCAC | AGTTAACAAC | ATCATTCAAG | TGAAGAAAAC | 600  |
| TGACTTTGAA | GTATTTGATG | CTTTAATGGT | GGATTCCCAA | AAGTGCTCCG | 650  |
| ATGTATCTGA | GCTGTCCAGT | TCCCCCCCAG | GACCCTACCA | CCAGGAGCCT | 700  |
| TACATCTCTA | AACCAGAGGA | GCGGTTCAAG | GCCCCGCCCC | TCCTCCCGCC | 750  |
| TCACCTGCTG | CAGGTCATCT | TGAACAAGGA | CACGGGCATC | TCTTGTGATC | 800  |
| CAGCGCTGCT | TCCGGAGCCC | AACCACGTCA | TGCTGAACCA | CCTCTATGCA | 850  |
| CTCTCTATCA | AGGATGGAGT | GATGGTGCTC | AGTGCGACCC | ATCGGTACAA | 900  |
| GAAAAAGTAC | GTCACCACCC | TCCTCTACAA | GCCCATATGA | GAGGATGAGC | 950  |
| CAGCCGTGGG | CCACGGGACA | GCAGGCGGGA | GCCGCTGGGC | TCTCCGTGTG | 1000 |
| CATGCGCATC | CTCACTCCGG | GACATCTCAC | CCCCACATAG | TCCTCCTTGA | 1050 |
| AGGTCTGTCC | AGGCACAGCC | AGAAATCGGA | TGGACGGCAG | ACCGTGGTCC | 1100 |
| CAGCACCGCA | GGCAGTGCGC | CAGGCTCTAG | TGCTCTAAGC | ATCATCCCTC | 1150 |
| TGCTGGCCCG | AGATGTCTAC | AGCCAGACCT | GAATGCTGGT | TCCTGCTAGA | 1200 |
| AAACCTAGGA | CAGGAAGTGA | AGTCACCAAA | GCCCTCATCA | TCCCTGCTGA | 1250 |
| AGCCTGGCTT | GGAAGAAAGC | AGTGCTCGGT | CTTGCTGTGC | CTTCCGAATC | 1300 |
| ACAGCAGTAG | ATTGTAGACT | CCATGGAATT | TCAGTGTCCA | ATTTCCAGAT | 1350 |
| GCAGCTTCGC | AATCGATTCC | TGACACTGTG | CACTGAGACC | TTCTTAACCA | 1400 |

- 52 -

```

GAGTGGCTGG CTGTCCACTC TCACCTTAAGG CAATAAGTCA CCAGGACGAG 1450
ACTATAGGTC ATGTGACTAC TGAGCAATAA TCGTTCTCAN ACAGACATCA 1500
GAAACCACTG CCATTTCTCC ATCAAGCCAG ACGATCCTGA GGA CTGACCA 1550
CCATGGGAGG TTGTCCACCT TATTTTCA GTT GCAGTGTG G CCATGTTACC 1600
GTGACAACCT GGTCGAAGTG CCCGCCCTCT TTTTAGTTCT AGCACGTGCT 1650
ACTCAGCTGG GGGCCGTGTC TCCAGTGAGC AGAGAGTGTA CACGGTGGTT 1700
ACTATTGCCT GATCCTAAGA GAGCTTGGCA CCCTGCGGCA GACTGCTAGG 1750
TTCCAGCAGG GTTGGCACGA GTGAACCTAT GTGTGCTCAG TGTGATTTCC 1800
ACAGTGATGT CACAGACGTG CCCATTGGTA CAGGCTCCTG TCACCTGTCA 1850
GCATAGGTAG GCACAAGCTC TGTGGTGTCC GCTATTTGGT TAAACCTGAG 1900
TTTTGGGTAC CTTTGTGTTAC TGTTTTCAAA ACACGGACTT GCTGTCATCT 1950
TGATGTACAA GTTTCAATAA AGCTTTGG 1978

```

(2) INFORMATION FOR SEQ ID NO: 62:

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 270

(B) TYPE: Amino Acid

(D) TOPOLOGY: Linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 62:

```

MET GLY ASN THR SER SER GLU ARG ALA ALA LEU GLU ARG GLN ALA
1      5      10      15
GLY HIS LYS THR PRO ARG ARG ASP SER SER GLU GLY THR LYS ASP
20     25     30
GLY ASP ARG PRO LYS ILE LEU MET ASP SER PRO GLU ASP ALA ASP
35     40     45
ILE PHE HIS THR GLU GLU MET LYS ALA PRO GLU LYS GLU GLU PHE
50     55     60
LEU ALA TRP GLN HIS ASP LEU GLU VAL ASN GLU LYS ALA PRO ALA
65     70     75
GLN ALA ARG PRO THR VAL PHE ARG TRP THR GLY GLY GLY LYS GLU
80     85     90
VAL TYR LEU SER GLY SER PHE ASN ASN TRP SER LYS LEU PRO LEU
95     100    105
THR ARG SER GLN ASN ASN PHE VAL ALA ILE LEU ASP LEU PRO GLU
110    115    120
GLY GLU HIS GLN TYR LYS PHE PHE VAL ASP GLY GLN TRP THR HIS
125    130    135
ASP PRO SER GLU PRO ILE VAL THR SER GLN LEU GLY THR VAL ASN
140    145    150
ASN ILE ILE GLN VAL LYS LYS THR ASP PHE GLU VAL PHE ASP ALA
155    160    165
LEU MET VAL ASP SER GLN LYS CYS SER ASP VAL SER GLU LEU SER
170    175    180
SER SER PRO PRO GLY PRO TYR HIS GLN GLU PRO TYR ILE SER LYS
185    190    195
PRO GLU GLU ARG PHE LYS ALA PRO PRO ILE LEU PRO PRO HIS LEU
200    205    210
LEU GLN VAL ILE LEU ASN LYS ASP THR GLY ILE SER CYS ASP PRO
215    220    225
ALA LEU LEU PRO GLU PRO ASN HIS VAL MET LEU ASN HIS LEU TYR

```

- 53 -

|                 |                     |                     |     |  |     |
|-----------------|---------------------|---------------------|-----|--|-----|
|                 | 230                 |                     | 235 |  | 240 |
| ALA LEU SER ILE | LYS ASP GLY VAL MET | VAL LEU SER ALA THR | HIS |  |     |
|                 | 245                 |                     | 250 |  | 255 |
| ARG TYR LYS LYS | LYS TYR VAL THR THR | LEU LEU TYR LYS PRO | ILE |  |     |
|                 | 260                 |                     | 265 |  | 270 |

(2) INFORMATION FOR SEQ ID NO: 63:

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 1576

(B) TYPE: Nucleic acid

(C) STRANDEDNESS: Single

(D) TOPOLOGY: Linear

(iv) ANTI-SENSE: No

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 63:

```

GCGCCCTTAA AGATGGTGAG GGGGCTATGC TCTGAGTAGA AGGTGGTGAC 50
CTCCAGGAGC GGTGGGATGA TGAGGGCCCG GGCGCCTCTT GCAATGGAGA 100
CGGTCAATTC TTCAGATAGC TCCCCAGCTG TGGAAAATGA GCATCCTCAA 150
GAGACCCCAG AATCCAACAA TAGCGTGTAT ACTTCCTTCA TGAAGTCTCA 200
TCGCTGCTAT GACCTGATTC CCACAAGCTC CAAATTGGTT GTATTTGATA 250
CGTCCCTGCA GGTGAAGAAA GCTTTTTTTG CTTTGGTGAC TAACGGTGTA 300
CGAGCTGCCC CTTTATGGGA TAGTAAGAAG CAAAGTTTTG TGGGCATGCT 350
GACCATCACT GATTTTCATCA ATATCCTGCA CCGCTACTAT AAATCAGCGT 400
TGGTACAGAT CTATGAGCTA GAAGAACACA AGATAGAAAC TTGGAGAGAG 450
GTGTATCTCC AGGACTCCTT TAAACCGCTT GTCTGCATTT CTCCTAATGC 500
CAGCTTGTTT GATGCTGTCT CTTCATTAA TCGCAACAAG ATCCACAGGC 550
TGCCAGTTAT TGACCCAGAA TCAGGCAATA CTTTGTACAT CCTCACCCAC 600
AAGCGCATTG TGAAGTTCCT CAAATTGTTT ATCACTGAGT TCCCCAAGCC 650
AGAGTTCATG TCCAAGTCTC TGGAAGAGCT ACAGATTGGC ACCTATGCCA 700
ATTTGCTATG GGTTCGCACT ACCACCCCGG TCTATGTGGC TCTGGGGATT 750
TTTGTACAGC ATCGAGTCTC AGCCCTGCCA GTGGTGGATG AGAAGGGGCG 800
TGTGGTGGAC ATCTACTCCA AGTTTGATGT TATCAATCTG GCAGCAGAAA 850
AGACCTACAA CAACCTAGAT GTATATGTGA CTAAAGCCTT GCAACATCGA 900
TCACATTACT TTGAGGGTGT TCTCAAGTGC TACCTGCATG AGACTCTGGA 950
GACCATCATC AACAGGCTAG TGGAAGCAGA GGTTCAACGA CTTGTAGTGG 1000
TGGATGAAAA TGATGTGGTC AAGGGAATTG TATCACTGTC TGACATCCTG 1050
CAGGCCCTGG TGCTCACAGG TGGAGAGAAG AAGCCCTGAG CTGGGGGAAG 1100
GGTCATGCAG CACCAGGGGA TATGCCCAAC TCACTGCCTG CTGGAAGCTC 1150
TGTGGGAATC AGATGAAACT TGAGGGAATT GTGACTCTGT TCCCTGTTCA 1200
GGGTCCCCTG CCCTTCTATC TGGGAGCTAG GGAAGGTATG GGGGAGGAAA 1250
GAGAATGGAT TTATAGCTAC CCTTACCCTC ACACATACAC TTGAAAAAAC 1300
TTTCAGCCTA GCCAGTTCTA GCCCCTGTCC TCTTAGATAT ATCCCCCTTT 1350
CTGGGTGAAC TATAGGCTCT GTGCCTCTCA GACAAATTCT GATCTCTAAG 1400
AGATCCCCAG ACCTCACTTG CCTCTGCCTC CATCTTGGCC CTGATTCAAC 1450
CCTAAGATAA TAGCACAACA AAATTCTTCA TAAAGATATT TTTATTCCAC 1500
TGTTCCGTGC TATATGGAGG AGGCCAAGTC CATTTAGTGA CATTTCTTCC 1550
CATAATGTGA GTGGGGAGGA TTGTGG 1576

```

(2) INFORMATION FOR SEQ ID NO: 64:

- 54 -

## (i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 331

(B) TYPE: Amino Acid

(D) TOPOLOGY: Linear

## (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 64:

|     |     |     |     |     |     |     |     |     |     |     |     |     |     |     |  |
|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|--|
| MET | GLU | THR | VAL | ILE | SER | SER | ASP | SER | SER | PRO | ALA | VAL | GLU | ASN |  |
| 1   |     |     |     | 5   |     |     |     |     | 10  |     |     |     |     | 15  |  |
| GLU | HIS | PRO | GLN | GLU | THR | PRO | GLU | SER | ASN | ASN | SER | VAL | TYR | THR |  |
|     |     |     |     | 20  |     |     |     |     | 25  |     |     |     |     | 30  |  |
| SER | PHE | MET | LYS | SER | HIS | ARG | CYS | TYR | ASP | LEU | ILE | PRO | THR | SER |  |
|     |     |     |     | 35  |     |     |     |     | 40  |     |     |     |     | 45  |  |
| SER | LYS | LEU | VAL | VAL | PHE | ASP | THR | SER | LEU | GLN | VAL | LYS | LYS | ALA |  |
|     |     |     |     | 50  |     |     |     |     | 55  |     |     |     |     | 60  |  |
| PHE | PHE | ALA | LEU | VAL | THR | ASN | GLY | VAL | ARG | ALA | ALA | PRO | LEU | TRP |  |
|     |     |     |     | 65  |     |     |     |     | 70  |     |     |     |     | 75  |  |
| ASP | SER | LYS | LYS | GLN | SER | PHE | VAL | GLY | MET | LEU | THR | ILE | THR | ASP |  |
|     |     |     |     | 80  |     |     |     |     | 85  |     |     |     |     | 90  |  |
| PHE | ILE | ASN | ILE | LEU | HIS | ARG | TYR | TYR | LYS | SER | ALA | LEU | VAL | GLN |  |
|     |     |     |     | 95  |     |     |     |     | 100 |     |     |     |     | 105 |  |
| ILE | TYR | GLU | LEU | GLU | GLU | HIS | LYS | ILE | GLU | THR | TRP | ARG | GLU | VAL |  |
|     |     |     |     | 110 |     |     |     |     | 115 |     |     |     |     | 120 |  |
| TYR | LEU | GLN | ASP | SER | PHE | LYS | PRO | LEU | VAL | CYS | ILE | SER | PRO | ASN |  |
|     |     |     |     | 125 |     |     |     |     | 130 |     |     |     |     | 135 |  |
| ALA | SER | LEU | PHE | ASP | ALA | VAL | SER | SER | LEU | ILE | ARG | ASN | LYS | ILE |  |
|     |     |     |     | 140 |     |     |     |     | 145 |     |     |     |     | 150 |  |
| HIS | ARG | LEU | PRO | VAL | ILE | ASP | PRO | GLU | SER | GLY | ASN | THR | LEU | TYR |  |
|     |     |     |     | 155 |     |     |     |     | 160 |     |     |     |     | 165 |  |
| ILE | LEU | THR | HIS | LYS | ARG | ILE | LEU | LYS | PHE | LEU | LYS | LEU | PHE | ILE |  |
|     |     |     |     | 170 |     |     |     |     | 175 |     |     |     |     | 180 |  |
| THR | GLU | PHE | PRO | LYS | PRO | GLU | PHE | MET | SER | LYS | SER | LEU | GLU | GLU |  |
|     |     |     |     | 185 |     |     |     |     | 190 |     |     |     |     | 195 |  |
| LEU | GLN | ILE | GLY | THR | TYR | ALA | ASN | ILE | ALA | MET | VAL | ARG | THR | THR |  |
|     |     |     |     | 200 |     |     |     |     | 205 |     |     |     |     | 210 |  |
| THR | PRO | VAL | TYR | VAL | ALA | LEU | GLY | ILE | PHE | VAL | GLN | HIS | ARG | VAL |  |
|     |     |     |     | 215 |     |     |     |     | 220 |     |     |     |     | 225 |  |
| SER | ALA | LEU | PRO | VAL | VAL | ASP | GLU | LYS | GLY | ARG | VAL | VAL | ASP | ILE |  |
|     |     |     |     | 230 |     |     |     |     | 235 |     |     |     |     | 240 |  |
| TYR | SER | LYS | PHE | ASP | VAL | ILE | ASN | LEU | ALA | ALA | GLU | LYS | THR | TYR |  |
|     |     |     |     | 245 |     |     |     |     | 250 |     |     |     |     | 255 |  |
| ASN | ASN | LEU | ASP | VAL | SER | VAL | THR | LYS | ALA | LEU | GLN | HIS | ARG | SER |  |
|     |     |     |     | 260 |     |     |     |     | 265 |     |     |     |     | 270 |  |
| HIS | TYR | PHE | GLU | GLY | VAL | LEU | LYS | CYS | TYR | LEU | HIS | GLU | THR | LEU |  |
|     |     |     |     | 275 |     |     |     |     | 280 |     |     |     |     | 285 |  |
| GLU | THR | ILE | ILE | ASN | ARG | LEU | VAL | GLU | ALA | GLU | VAL | HIS | ARG | LEU |  |
|     |     |     |     | 290 |     |     |     |     | 295 |     |     |     |     | 300 |  |
| VAL | VAL | VAL | ASP | GLU | ASN | ASP | VAL | VAL | LYS | GLY | ILE | VAL | SER | LEU |  |
|     |     |     |     | 305 |     |     |     |     | 310 |     |     |     |     | 315 |  |
| SER | ASP | ILE | LEU | GLN | ALA | LEU | VAL | LEU | THR | GLY | GLY | GLU | LYS | LYS |  |
|     |     |     |     | 320 |     |     |     |     | 325 |     |     |     |     | 330 |  |
| PRO |     |     |     |     |     |     |     |     |     |     |     |     |     |     |  |



**What is Claimed is:**

1. A nucleic acid sequence encoding mammalian AMPK  $\alpha_1$ .
2. The nucleic acid sequence of claim 1 comprising SEQ ID NO: 44.
3. A vector comprising a nucleic acid sequence of claim 1.
4. A host cell comprising a vector of claim 3.
5. A recombinant polypeptide encoded by the nucleic acid sequence of claim 1.
6. A method of producing mammalian AMPK  $\alpha_1$  comprising:
  - (a) culturing cells of claim 4 under conditions which allow expression of the nucleic acid sequence encoding mammalian AMPK  $\alpha_1$ ; and
  - (b) recovering the expressed AMPK  $\alpha_1$  from the cell.
7. An oligonucleotide probe comprising at least 10 nucleotides, said oligonucleotide probe being capable of selectively hybridizing to a nucleic acid sequence of claim 1.
8. A substantially purified polypeptide or biologically active fragment thereof encoded by a nucleic acid sequence of claim 1.
9. An antibody capable of binding selectively to a polypeptide of claim 8.
10. A nucleic acid sequence encoding mammalian AMPK  $\beta$ , said nucleic acid sequence comprising SEQ ID NO: 61.
11. A vector comprising the nucleic acid sequence of claim 10.

12. A host cell comprising a vector of claim 11.
13. A recombinant polypeptide encoded by the nucleic acid sequence of claim 10.
14. A method of producing mammalian AMPK  $\beta$  comprising:  
5 (a) culturing cells of claim 12 under conditions which allow expression of the nucleic acid sequence encoding AMPK  $\beta$ ; and  
(b) recovering the expressed AMPK  $\beta$ .
15. A substantially purified polypeptide comprising an  
10 amino acid sequence of SEQ ID NO: 62.
16. A nucleic acid sequence encoding mammalian AMPK  $\gamma$ , said nucleic acid sequence comprising SEQ ID NO: 63.
17. A vector comprising the nucleic acid sequence of claim 16.
- 15 18. A host cell comprising a vector of claim 17.
19. A recombinant polypeptide encoded by the nucleic acid sequence of claim 16.
20. A method of producing mammalian AMPK  $\gamma$  comprising:  
20 (a) culturing cells of claim 18 under conditions which allow expression of the nucleic acid sequence encoding AMPK  $\gamma$ ; and  
(b) recovering the expressed AMPK  $\gamma$ .
21. A substantially purified polypeptide comprising an amino acid sequence of SEQ ID NO: 64.

## INTERNATIONAL SEARCH REPORT

International application No.  
PCT/US97/00270

## A. CLASSIFICATION OF SUBJECT MATTER

IPC(6) : C07K 2/00, 14/47, 16/18; C12N 5/10, 15/09, 15/11, 15/12, 15/63, 15/70, 15/74, 15/79

US CL : 435/69.1, 320.1, 325, 252.3, 254.11; 530/300, 350, 387.1, 412; 536/23.5, 24.31

According to International Patent Classification (IPC) or to both national classification and IPC

## B. FIELDS SEARCHED

Minimum documentation searched (classification system followed by classification symbols)

U.S. : 435/69.1, 320.1, 325, 252.3, 254.11; 530/300, 350, 387.1, 412; 536/23.5, 24.31

Documentation searched other than minimum documentation to the extent that such documents are included in the fields searched

Electronic data base consulted during the international search (name of data base and, where practicable, search terms used)

APS, STN (MEDLINE, INPADOC, EMBASE, CAPLUS, WPIDS)

search terms: AMP, protein kinase?, alpha?, beta?, gamma?, 5' AMP, AMPK, stapleton d7/au, mitchelhill k7/au, witters l7/au

## C. DOCUMENTS CONSIDERED TO BE RELEVANT

| Category*         | Citation of document, with indication, where appropriate, of the relevant passages  | Relevant to claim No.    |
|-------------------|---|--------------------------|
| X<br>----, P<br>Y | WOODS, A. et al. Characterization of AMP-activated protein kinase beta and gamma subunits. J. Biol. Chem. 26 April 1996. Vol. 271, No. 5, pages 10282-10290, especially page 10283 and Fig. 1-2.  | 10-15<br>-----<br>16-21  |
| X<br>---<br>Y     | CARLING D. et al. Mammalian AMP-activated protein kinase is homologous to yeast and plant protein kinases involved in the regulation of carbon metabolism. J. Biol. Chem. 15 April 1994. Vol. 269, No. 15, pages 11442-11448, especially page 11444 and Fig. 3. | 1-3, 7-9<br>-----<br>5,6 |
| X<br>---, P<br>Y  | STAPLETON D. et al. Mammalian AMP-activated protein kinase subfamily. J. Biol. Chem. 12 January 1996. Vol. 271, No. 2, pages 611-614, especially page 612.  | 1,2,7-9<br>-----<br>3-6  |

☒ Further documents are listed in the continuation of Box C. ☐ See patent family annex.

|  |     |  |
|--|-----|--|
| * Special categories of cited documents:   | * T | later document published after the international filing date or priority date and not in conflict with the application but cited to understand the principle or theory underlying the invention  |
| * A* document defining the general state of the art which is not considered to be of particular relevance  | * X | document of particular relevance; the claimed invention cannot be considered novel or cannot be considered to involve an inventive step when the document is taken alone   |
| * E* earlier document published on or after the international filing date  | * Y | document of particular relevance; the claimed invention cannot be considered to involve an inventive step when the document is combined with one or more other such documents, such combination being obvious to a person skilled in the art |
| * L* document which may throw doubts on priority claim(s) or which is cited to establish the publication date of another citation or other special reason (as specified) | * Z | document member of the same patent family  |
| * O* document referring to an oral disclosure, use, exhibition or other means  |     |  |
| * P* document published prior to the international filing date but later than the priority date claimed  |     |  |

Date of the actual completion of the international search

07 APRIL 1997

Date of mailing of the international search report

25 APR 1997

Name and mailing address of the ISA/US  
Commissioner of Patents and Trademarks  
Box PCT  
Washington, D.C. 20231

Facsimile No. (703) 305-3230

Authorized officer

CLAIRE M. KAUFMAN

Telephone No. (703) 308-0196

## INTERNATIONAL SEARCH REPORT

International application No.  
PCT/US97/06270

## C (Continuation). DOCUMENTS CONSIDERED TO BE RELEVANT

| Category* | Citation of document, with indication, where appropriate, of the relevant passages  | Relevant to claim No. |
|-----------|---|-----------------------|
| X         | BERI R.K. et al. Molecular cloning, expression and chromosomal localisation of human AMP-activated protein kinase. FEBS Lett. 1994. Vol. 356, pages 117-121.  | 1,7,9                 |
| X         | WO 94/28116 A (ZENECA LIMITED) 08 December 1994, see entire document.   | 1,3-9                 |
| X         | STAPLETON D. et al. Mammalian 5'-AMP-activated protein kinase non-catalytic subunits are homologs of proteins that interact with yeast Snf1 protein kinase. J. Biol. Chem. 25 November 1994. Vol. 269, No. 47, pages 29343-29346, especially bottom of 1st col. | 1,15                  |
| X         | MITCHELHILL K.I. et al. Mammalian AMP-activated protein kinase shares structural and functional homology with the catalytic domain of yeast Snf1 protein kinase. J. Biol. Chem. 28 January 1994. Vol. 269, No. 4, pages 2361-2364, especially Fig. 1.           | 8                     |
| X         | YANG X. et al. A family of proteins containing a conserved domain that mediates interaction with the yeast SNF1 protein kinase complex. EMBO J. 1994. Vol. 13, No. 24, pages 5878-5886, especially Fig. 2.  | 15                    |
| X         | GAO G. et al. Catalytic subunits of the porcine and rat 5'-AMP-activated protein kinase are members of the SNF1 protein kinase family. Biochim. Biophys. Acta. 1995. Vol. 1266, pages 73-82, especially pages 74-76.  | 1-3,7                 |
| A         | CELENZA J.L. et al. Molecular analysis of the SNF4 gene of <i>Saccharomyces cerevisiae</i> : Evidence for physical association of the SNF4 protein with the SNF1 protein kinase. Mol. Cell. Biol. November 1989. Vol. 9, No. 11, pages 5045-5054.               | 1-9,16-21             |
| A, P      | PIOSIK P.A. et al. Carpine homologue of rodent 5'-AMP-activated protein kinase subunit and yeast SNF4/CAT3 is down-regulated by thyroid hormone. Mol. Brain Res. 1996. Vol. 40, pages 240-253.  | 16-21                 |
| A         | AGUAN K. et al. Characterization and chromosomal localization of the human homologue of a rat AMP-activated protein kinase-encoding gene: a major regulator of lipid metabolism in mammals. Gene. 1994. Vol. 149, pages 345-350.                                | 1-7                   |

## INTERNATIONAL SEARCH REPORT

International application No.  
PCT/US97/00270

## C (Continuation). DOCUMENTS CONSIDERED TO BE RELEVANT

| Category* | Citation of document, with indication, where appropriate, of the relevant passages  | Relevant to claim No. |
|-----------|---|-----------------------|
| A         | DALE S. et al. Similar substrate recognition motifs for mammalian AMP-activated protein kinase, higher plant HMG-CoA reductase kinase-A, yeast SNF1, and mammalian calmodulin-dependent protein kinase I. FEBS Lett. 1995. Vol. 361, pages 191-195. | 1-8                   |
| A, P      | DYCK J.R.B. et al. Regulation of 5'-AMP-activated protein kinase activity by the noncatalytic beta and gamma subunits. J. Biol. Chem. 26 July 1996. Vol. 271, No. 30, pages 17798-17803.  | 10-21                 |
| A         | VERHOEVEN A.J.M. et al. The AMP-activated protein kinase gene is highly expressed in rat skeletal muscle. Eur. J. Biochem. 1995. Vol. 228, pages 236-243.   | 1-7                   |
| A, P      | MICHELL B.J. et al. Isoform-specific purification and substrate specificity of the 5'-AMP-activated protein kinase. J. Biol. Chem. 08 November 1996. Vol. 271, No. 45, pages 28445-28450.   | 1-21                  |

# INTERNATIONAL SEARCH REPORT

International application No.  
PCT/US97/00270

## Box I Observations where certain claims were found unsearchable (Continuation of item 1 of first sheet)

This international report has not been established in respect of certain claims under Article 17(2)(a) for the following reasons:

1. ☐ Claims Nos.:  
because they relate to subject matter not required to be searched by this Authority, namely:
2. ☐ Claims Nos.:  
because they relate to parts of the international application that do not comply with the prescribed requirements to such an extent that no meaningful international search can be carried out, specifically:
3. ☐ Claims Nos.:  
because they are dependent claims and are not drafted in accordance with the second and third sentences of Rule 6.4(a).

## Box II Observations where unity of invention is lacking (Continuation of item 2 of first sheet)

This International Searching Authority found multiple inventions in this international application, as follows:

Please See Extra Sheet.

1. ☒ As all required additional search fees were timely paid by the applicant, this international search report covers all searchable claims.
2. ☐ As all searchable claims could be searched without effort justifying an additional fee, this Authority did not invite payment of any additional fee.
3. ☐ As only some of the required additional search fees were timely paid by the applicant, this international search report covers only those claims for which fees were paid, specifically claims Nos.:
4. ☐ No required additional search fees were timely paid by the applicant. Consequently, this international search report is restricted to the invention first mentioned in the claims; it is covered by claims Nos.:

Remark on Protest

☐

The additional search fees were accompanied by the applicant's protest.

☒

No protest accompanied the payment of additional search fees.

## INTERNATIONAL SEARCH REPORT

International application No.  
PCT/US97/00270

### BOX II. OBSERVATIONS WHERE UNITY OF INVENTION WAS LACKING

This ISA found multiple inventions as follows:

This application contains the following inventions or groups of inventions which are not so linked as to form a single inventive concept under PCT Rule 13.1. In order for all inventions to be searched, the appropriate additional search fees must be paid.

Group I, claims 1-8, drawn to nucleic acid encoding AMPK alpha1 polypeptide, oligonucleotide, vector, host cell, polypeptide, and method of producing polypeptide.

Group II, claim 9, drawn to antibody.

Group III, claims 10-15, drawn to nucleic acid encoding AMPK beta polypeptide, oligonucleotide, vector, host cell, polypeptide, and method of producing polypeptide.

Group IV, claims 16-21, drawn to nucleic acid encoding AMPK gamma polypeptide, oligonucleotide, vector, host cell, polypeptide, and method of producing polypeptide.

The inventions listed as Groups I-IV do not relate to a single inventive concept under PCT Rule 13.1 because, under PCT Rule 13.2, they lack the same or corresponding special technical features for the following reasons: The nucleic acid, polypeptide, vector, host cell, and method of Groups I, III, and IV do not share a special technical with each other because each group relates to different AMPK subunits (alpha1, beta, and gamma, respectively) that do not share structure or function. Additionally, Groups I, III, and IV do not share a special technical feature with the antibody of Group II because that antibody does not share structure or function with the other products and cannot be made by or used in any of the claimed methods.